I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17472

Small-Scale Farmers' Willingness to Pay for Information: A Comparison of Individual Purchase Decisions with Contributions to a Club Good

Michael Grimm
Nathalie Luck
Alia Bihrajihant Raya
Udit Sawhney

NOVEMBER 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17472

Small-Scale Farmers' Willingness to Pay for Information: A Comparison of Individual Purchase Decisions with Contributions to a Club Good

Michael Grimm
University of Passau, RWI Research Network and IZA

Nathalie Luck
University of Passau and TU Munich

Alia Bihrajihant Raya
Gadjah Mada University, Yokyakarta

Udit Sawhney
University of Passau

NOVEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Small-Scale Farmers' Willingness to Pay for Information: A Comparison of Individual Purchase Decisions with Contributions to a Club Good*

Soil tests provide information that can help farmers to reduce costs, increase yields and profits, and contribute to sustainable soil health, yet they are rarely used. In this study, we elicit small-scale farmers' willingness-to-pay (WTP) for rapid, low-cost soil tests using incentive-compatible auctions. Additionally, we test whether randomized participation in a sustainable soil management training can increase farmers' WTP. Furthermore, we elicit an alternative WTP by measuring the willingness to contribute to the costs of a soil test kit when farmer groups are offered kits containing 50 tests along with a training session that enables them to carry out the soil tests independently. Free riding is possible in this setting, and contributions will depend on social preferences and beliefs about the contributions of others. Our study shows that the WTP for soil information is considerable. Although we find some evidence for free riding, this does not significantly affect the WTP between the individual and the group auction. Our experiment demonstrates that integrating soil tests with existing extension services could be relatively straightforward. Subsidies can be justified by the potential environmental benefits.

JEL Classification: D71, H23, O13, O33, Q16, Q57

Keywords: technology adoption, soil tests, information constraints, willingness-to-pay, sustainability, public good games

Corresponding author:
Michael Grimm
School of Business Administration, Economics, and Information Systems
University of Passau
Innstrasse 27
94032 Passau
Germany
E-mail: michael.grimm@uni-passau.de

* Funding from the German Federal Environmental Foundation (DBU) is gratefully acknowledged (Grant # 36088/01). We thank Mardha Tilla Septiani and Nurul Husna Primasari for their valuable support in the field. Moreover, we thank participants at GDE 2024 in Hannover, at the summer BGPE workshop 2024 and at the TUM conference on "The Economics of the Food System: Transition toward Sustainability" for valuable comments and suggestions. The experiment has been pre-registered in the AER RCT Registry. Trial ID: AEARCTR-0011773.

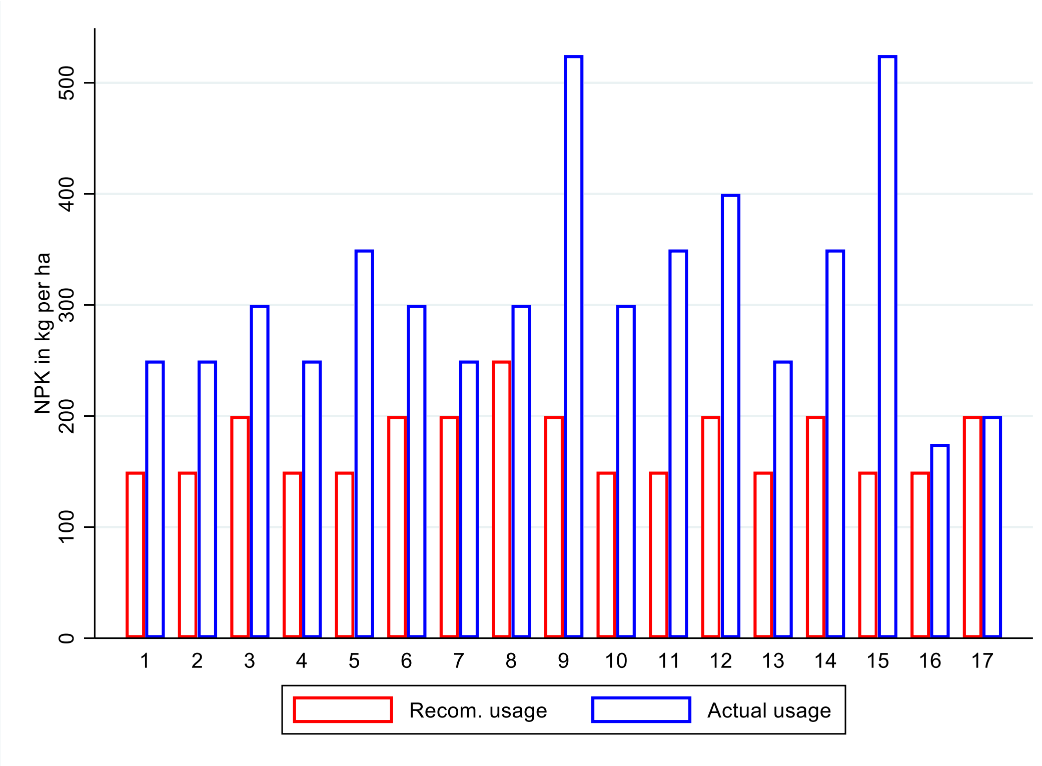
1. Introduction

The over-application of fertilizer is very common, especially in countries where fertilizer is heavily subsidized, such as Indonesia and Malawi. This over-application reduces farmers' profits and harms the environment (Lee, 2005), and degraded soils reduce yields in the long term (Hazell, 2009; IFAD, 2013; Lai, 2017). In other countries, a substantial share of farmers either uses insufficient fertilizer or none at all, which also reduces their yields and profits (Langyintuo, 2020). The literature has identified various reasons for this, including information constraints, behavioral biases, weak beliefs, affordability, and information asymmetries about the market for inputs, among others (see e.g., Duflo et al., 2011; Bold et al., 2017; Beg et al., 2024). This paper focuses on situations where smallholders use too much fertilizer, do not always use the right mix, and apply it at the wrong time. In such cases, soil tests can provide information that allows farmers to increase their yields and profits and manage their soils more sustainably (Beaman et al., 2013; Islam and Beg, 2021). Yet soil tests are rarely used, also because their costs typically exceed what extension services can cover. This raises the question of how soil tests could be provided to farmers in a way that ensures adoption and is at least partially cost covering. If farmers are willing to contribute to the cost of soil tests, they could be offered more widely.

Hence, the objectives of this paper are threefold. First, we elicit small-scale farmers' willingness-to-pay (WTP) for rapid, low-cost paddy soil tests using incentive-compatible auctions based on the Becker-DeGroot-Marschak (BDM) approach. Here, farmers are offered the service of having one (or several) plot(s) tested on the spot by an extension worker who then provides individualized fertilizer recommendations (the 'service arm' hereafter). Since about a third of our participants were randomly assigned to a one-day training on soil fertility management prior to the experiment, we can also causally determine whether information about the benefits of appropriate input management increases the WTP for soil tests. Second, we measure the willingness to contribute to the costs of a soil test kit when groups of farmers are offered soil test kits containing materials for up to 50 tests and a training session that enables them to carry out the soil tests themselves (the 'club good arm' hereafter). Free riding is possible in this setting, and contributions will depend on social preferences and beliefs about the contributions of others. Third, we compare demand in both settings to see which mode of provision leads to a higher uptake of the technology. We make this comparison for varying levels of subsidies.

Our experiment was implemented in Indonesia, in 45 villages in the province of Yogyakarta. As in the rest of Java, rice is by far the most important crop in Yogyakarta. In Indonesia, fertilizer has been heavily subsidized since the Green Revolution, and the resulting over-application of chemical inputs has been identified as one of the major causes of soil degradation (Simatupang and Timmer, 2008; Mariyono 2014; Alta et al., 2021). Nitrogen pollution and soil degradation can occur not only from fertilizer over-application but also from the incorrect timing of its application (Norton et al., 2015). In Indonesia, the indiscriminate application of nitrogen-rich fertilizers has become a widespread phenomenon among rice farmers in recent years (Widowati et al., 2011; Mariyono et al., 2018; Sukayat et al., 2023). This over-application has made crops vulnerable to pests and diseases, causing plants to collapse (Stevens et al., 1999) and polluting ground water (Bijay-Singh and Craswell, 2021). A recent (non-randomized) demonstration experiment conducted by the Indonesian Soil Research Institute (ISRI[1]), a department of the Indonesian Ministry of Agriculture, with paddy rice farmers across 17 different sites in West Java revealed that farmers consistently applied more NPK fertilizer (nitrogen, phosphorus, and potassium) than recommended by a soil test (Irawan et al., 2021). On average, farmers used 55% more NPK fertilizer than the recommended amounts (Figure 1).[2]

Figure 1. NPK overuse in kg per ha at 17 different sites in West Java, Indonesia



Notes: Two plots were compared at each site. On one plot fertilizer was applied by the farmer as usual and on the other fertilizer was applied according to soil test recommendations.
Source: Own representation, based on data by Irawan et at. (2021).

In our sample for this present study, about 40% of the respondents also overapplied chemical nitrogen to their fields, which is also the major chemical component of urea.

The soil test we used was developed by the ISRI. The test provides information about nutrient availability in the soil and makes recommendations to address nutrient deficiencies. Results are available within 30 to 45 minutes, with the analysis conducted directly in the field, thus eliminating the need for a laboratory, which usually takes two weeks. The costs of one test amount to approximately USD 2.30,[3] very little compared to the potential benefits: ISRI's demonstration project suggested that optimal input use can increase yields by 0.5 to 1 ton per ha (6% to 13%), and net benefit – the value of production minus additional costs – by USD 137

[1] Now called Indonesian Soil and Fertilizer Standardization Institute (BPSI Tanah dan Pupuk). In this paper, we refer to the organization as "ISRI".
[2] The farmers in the sample used on average 16% less urea (as another source of nitrogen) than recommended, but this does not compensate for the much larger overuse of NPK fertilizer.
[3] Exchange rate used: 1 USD = 15,725 IDR.

to 256 per ha (Irawan et at., 2021). Given the average plot size in our sample of 0.17 ha, this would amount to USD 23 to 44 per plot.

The findings of our study are as follows. The willingness-to-pay for soil information is considerable, indicating that providing soil tests through extension workers at a subsidized rate is feasible. Farmers are, on average, willing to pay about USD 1, which is almost half of the cost covering price of the test (not counting the costs of performing the test or training). On a per test basis, both types of provisions (service and club good) result in comparable price bids. Although we find some evidence for free riding, this does not significantly affect the WTP between the service and the club good. A priori one-day training on soil fertility management does not seem to significantly affect the value attributed to fertilizer recommendations, but our results suggest that the training has at least a small positive effect on the total budget that farmers are willing to spend on soil tests.

With this study we contribute to the literature on technology adoption, specifically to the literature on the role of information constraints in technology adoption and input use (Foster and Rosenzweig, 1995; Conley and Udry, 2001, 2010; Jack, 2011; Krishnan and Patnam, 2014; Magruder, 2018; Tamim et al., 2020). We explore the extent to which small-scale farmers are willing to pay for information which they may perceive as more or less useful. Unlike many existing studies that measure stated WTP (see e.g., Kokoye et al., 2018), we elicit revealed WTP, similar to Berazneva et al. (2023) in Malawi and Fabregas et al. (2020) in Kenya. While Berazneva et al. (2023) also conducted a group experiment, in their setting participants collectively bid for a single test in their village, under the assumption that a test on one plot has value for all but that this value declines with perceived soil heterogeneity. They did not use a BDM auction but rather played a threshold public goods game. Another difference is that participants in their study were endowed with a cash amount before they made their bid. Such a windfall gain may result in a different WTP than if the bid has to be made from regular income, for example because participants may take more risks with a windfall gain (see e.g., Jing and Cheo, 2013). The authors randomly varied the plot in the village to be tested to explore the role of soil heterogeneity. They found that farmers contributed more when they perceived soils to be similar to their own plot but also free ride on others' contributions. Alternatively, Fabregas et al. (2020) used a lottery incentive system and the BDM auction to elicit the WTP for soil test results from local areas and local experimental plots. They concluded that the aggregate WTP for soil information in any given area exceeded the costs of generating and disseminating such information. In both cases, soil tests were performed at least partly in the lab, unlike our study where tests were conducted on-site.

Other studies have not measured the WTP for soil tests, but rather how soil test information alters the WTP for inputs. Fishman et al. (2019), for example, conducted a randomized controlled trial in the Indian state of Bihar to test a government program of targeted soil testing and customized fertilizer recommendations. They found no evidence that soil testing and targeted fertilizer recommendations affected fertilizer use or that farmers' WTP for micronutrients changed, suggesting that farmers did not attribute value to the information from the soil tests. They explored the role of confidence, finding that more confident farmers had a lower stated WTP for soil testing ex-ante and a lower responsiveness to recommended application rates ex-post. Low literacy levels and the perceived credibility of the information were further negatively associated with adherence to recommendations. Similarly, Murphy et al. (2020) conducted an experimental auction in Western Kenya to determine whether providing information and fertilizer recommendations from low-cost soil tests affected farmers' WTP for organic and inorganic agricultural inputs. In contrast to Fishman et al. (2019), they found that providing soil information significantly affected farmers' revealed WTP for inputs and that the benefits of better input choices likely offset the costs of the soil tests.

In summary, our experiment differs in three respects from the existing literature. First, we focus on individual soil test information rather than village-level information, thus addressing the potential problem of soil heterogeneity and hence the limited value of information coming from plots other than the farmers' own. Second, we use low-cost soil tests that can be carried out on-site and, with some training, by the farmers themselves. Third, we compare the provision of a soil test as an individual service with its provision as a test kit given to farmer groups. Additionally, beyond contributing to the literature on technology adoption and information constraints, we also add to the literature that explores strategies to achieve both higher productivity and environmental sustainability (see e.g., Lee, 2005; Michler et al., 2019; Grimm and Luck, 2023).

The remainder of this paper is organized as follows. In Section 2, we briefly introduce the soil test used in this study. In Section 3, we explain our sampling and randomization procedures and provide a description of our sample. In Section 4, we describe the WTP experiments, provide a theoretical framework, and derive some hypotheses. In Section 5, we present our results, including revealed demand curves and an analysis of the two alternative provision channels. In Section 6 we discuss cost-effectiveness and the potential role of subsidies. In Section 7 we conclude.

2. The soil test

Participants in our study reveal their WTP for a rapid paddy rice soil test. The soil test was developed by ISRI, which is, as mentioned above, a department of the Indonesian Ministry of Agriculture. This test has been tested and validated both in the laboratory and in the field (Widowati et al., 2004; Irawan et al., 2021). It measures the levels of nitrogen (N), phosphorus (P), potassium (K), and pH in the soil, enabling farmers to choose the appropriate fertilizer and dosage. The test also provides general recommendations on the timing of application. It can be conducted directly in the field, with results available within 30 to 45 minutes, eliminating the need to send soil samples to a laboratory. The accuracy of the tests is relatively high, a validation study based on 146 soil samples comparing the results from the rapid soil test with those obtained in the lab yielded an average alignment for N, P, K, and Ph of nearly 80% (Widowati et al., 2004).

After receiving instructions and training, farmers can, in theory, conduct the test without expert support. A composite soil sample can represent three to five hectare of a relatively homogeneous stretch of paddy fields. However, if neighboring paddy fields are managed by different farmers with different fertilizer strategies, separate tests may be necessary. According to ISRI, the validity of test-based recommendations lasts about one to two years or three to six rice growing seasons (Irawan et al., 2021). Hence, it is a rapid and low-cost technology with a high scalability potential. If used correctly, it has been shown to significantly increase farmers' yields and profits while reducing fertilizer over-application (Irawan et al., 2021).

The tests are marketed as kits (PUTS), which comprise of test tubes and liquids for 50 soil tests, along with a bag and a user manual. A PUTS costs IDR 1.8 million (approximately USD 116).

Figure 2: The soil test kit

Source: ISRI.

In our experiment, we offered the tests in two alternative forms, either as a service where one or several tests were carried out by an extension officer who then also provided fertilizer recommendations based on the test results, or as an entire kit offered to a group of farmers with group training. After training, farmers were expected to conduct the up to 50 tests themselves. The group training incurred additional costs of IDR 250,000 (USD 15.90). A single test costs IDR 36,100 (USD 2.30) and requires services at a value of IDR 25,000 (USD 1.59).

3. Sampling, randomization, and sample description

For this study, we sampled 46 villages (sub-villages or Dusuns to be precise) from a larger sample of 69 villages that had previously been identified for related project, the "Sustainable Soil Management Project" or "SSM-project" hereafter. The 69 villages for the SSM-project were randomly selected using a multi-stage sampling design in the province of Yogyakarta (for details see Appendix A).

The SSM-project was designed as a randomized controlled trial (RCT). Villages were randomly assigned to one of three groups: (i) a group offered a one-day soil management training, (ii) a group offered a two-day soil management and soil testing training, and (iii) a control group. For our present study, we drew only from the group offered the soil management training (23 villages) and from the control group (23 villages). We did not include villages from the group that was also offered soil testing training as these farmer groups had already received a soil test kit as part of the training and hence asking them about their WTP was not meaningful.

Of the 46 villages we drew from the SSM-project, we initially randomized 28 villages into the club good experimental arm and 18 villages into the service good experimental arm. However, villages initially assigned to the club good arm where less than ten farmers attended the experiment were reassigned to the service arm.[4] Additionally, one village had to be dropped entirely as no one wanted to participate. Ultimately, we assigned 21 villages to the club good arm and 24 villages to the service arm. In the club good arm, nine villages had received the one-

[4] Below we test the robustness of our findings towards this re-assignment.

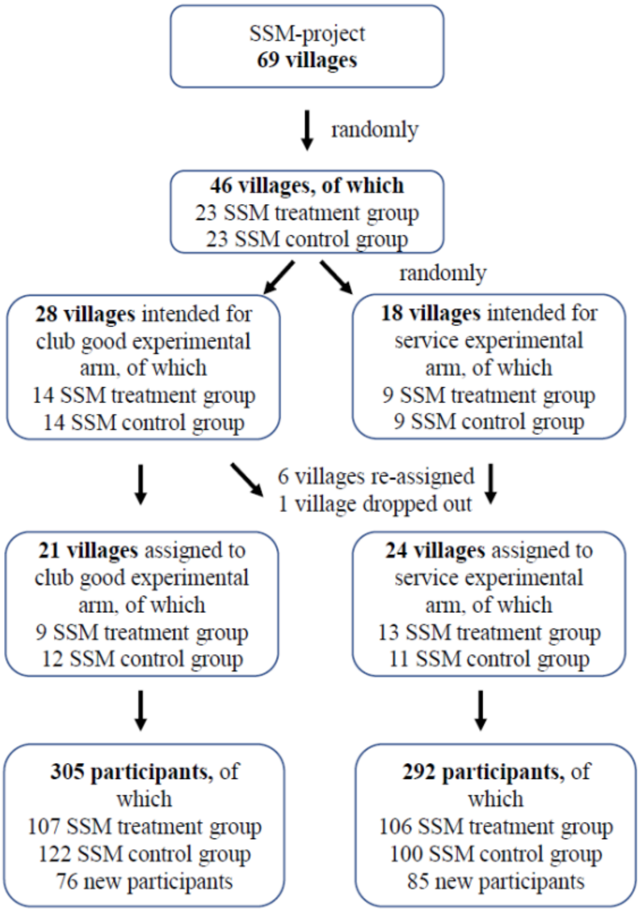
day soil management training prior to the WTP experiment, while 13 villages in the service arm had received this training.

In each village, we invited up to 16 participants to an information session about soil testing. The invited farmers were those who had participated in the baseline and endline surveys of the SSM-project. We did not expect all invited farmers to show up and did not deny attendance to farmers who were not part of the SSM-project and hence were not invited but showed an interest. In total, we gathered 597 participants – 292 in the service arm and 305 in the club good arm. Among these participants, 161 were new and had not been involved in any previous stages of the SSM-project. Almost all participants were smallholders cultivating an average land size of 0.17 ha. The main crop that is cultivated in the area is rice. Figure 3 summarizes the sampling procedure.

During the WTP experiment, we collected only limited data about the participants to avoid burdening them with another very time-consuming task, which might have drawn their attention away from the bidding process. Table 1 shows the characteristics which were covered. Participants were, on average, 53.2 years old, 22.4% had primary education or lower, and 59.0% had completed senior high school or higher. Furthermore, 35.7% of all participants had benefited from the training offered by the SSM-project and had been surveyed, while 37.2% had been in the control group of the SSM-project and also surveyed. About 27.0% were new additions to the sample and thus had not been surveyed or offered any training in the SSM-project. Table 1 shows that all characteristics are reasonably balanced across the service and club good arms.

We use the sub-sample of SSM-project participants and the data collected there to provide additional information on farming practices. Most respondents were the head of their household and hence likely to make financial decisions in their household. For about half of them, farming was the main activity, while most of the others reported owning non-farm enterprises or being wage laborers as their primary economic activity. Respondents cultivated an average of 3.4 plots, 61.4% used manure and more than 50% used organic fertilizers other than manure. As expected, virtually all farmers used inorganic fertilizers such as urea. Of the participants, 38.9% had already heard of soil tests (but not through the SSM-project) and 17% had conducted a soil test at least once, though not necessarily themselves. On average, per farmer, 37.0% of all rice plots were under sharecropping, 13.9% under fixed rent, and 44.5% of all rice plots were owned, i.e., many farmers cultivated several plots under different ownership and tenancy arrangements.

Figure 3. Sampling and participant flow



Source: Own representation.

Table 1: Sample description

	All particip.	Service	Club good	*t*- test (*p*-value)
WTP sample - all participants				
Age	53.2	53.7	52.8	0.290
Primary or lower (=1)	0.224	0.226	0.223	0.928
Junior high (=1)	0.186	0.188	0.184	0.882
Senior high or higher (=1)	0.590	0.586	0.593	0.846
No prior training, surveyed (=1)	0.372	0.342	0.400	0.146
Prior training, surveyed (=1)	0.357	0.363	0.351	0.756
New participant, no prior survey (=1)	0.270	0.291	0.249	0.249
Obs.	507	292	305	
WTP sample - without new participants				
Household head (=1)	0.867	0.894	0.843	n.a.
Main activity farming (=1)	0.526	0.490	0.559	n.a.
# plots cultivated by hh	3.347	3.369	3.328	n.a.
# rice plots cultivated by hh	2.400	2.534	2.279	n.a.
Used manure (=1)†	0.614	0.602	0.624	n.a.
Used any organic fertilizer, excl. manure (=1) †	0.550	0.531	0.568	n.a.
Used any inorganic fertilizer (=1)	0.993	0.986	1.000	n.a.
Heard of soil test (=1)	0.389	0.408	0.371	n.a.
Has already tested soil at least once (=1)	0.170	0.184	0.157	n.a.
Share of rice fields under sharecropping (=1)	0.370	0.383	0.357	n.a.
Share of rice fields under fixed rent (=1)	0.139	0.113	0.163	n.a.
Share of rice fields owned (=1)	0.445	0.469	0.423	n.a.
Obs.	436	207	229	

Notes: † Used manure and used organic fertilizer relate to rice plots only.

The one-day soil management training, which 169 of our respondents were offered within the SSM-project, was similar in structure and length to typical training sessions conducted by extension officers. It covered the importance of soil nutrient management, the role of organic matter in maintaining good soil structure, the use of leaf color charts, the principles of low external input sustainable agriculture (LEISA), and the use of digital resources for additional information on sustainable farming practices. The training also included a practical session on preparing organic inputs. The training program was jointly designed by our team, colleagues from the Agricultural Faculty of Universitas Gadjah Mada (UGM), and trainers from P4S. P4S are self-help agricultural and rural training groups that are owned and managed by farmers. They exist in most districts in Indonesia and receive financial resources from the local government. The training took place between July and September 2022. Our WTP experiment was implemented one year later, between July and September 2023. We further conducted a short follow-up survey in February 2024 which included questions on whether respondents discussed the price with others before making their bids and how many tests they had carried out since the experiment.

4. The willingness-to-pay experiment and hypotheses

4.1 The BMD auction and the experimental arms

To measure the WTP for soil tests, we used a Becker-DeGroot-Marschak auction (BDM). BDM is widely used in the literature and has the advantage of being incentive-compatible.[5] The principle of BDM is to offer a respondent a product and ask for a price bid. This bid is then compared to a price randomly drawn from a distribution determined by the researchers. If the bid is equal to or higher than the drawn price (the strike price), the respondent buys the product at the drawn price. If the bid is lower, the respondent cannot buy the product. The underlying assumption is that this procedure can reveal the true WTP of a rational consumer, as respondents who bid too low may miss out on buying a product at a price below their valuation, while bidding too high would force them to pay more than their valuation. Yet there is still a risk that respondents might gamble if they do not interpret the situation as a real purchase decision. Respondents may also struggle to provide a WTP if the product is entirely new to them and they cannot assess its value. In such cases, researchers sometimes provide a price range. The downside of providing a price range is that it can lead to anchoring at the lower and upper ends of the range (Grimm et al., 2020). In our study, we did not provide a price range but informed participants of the purchase price of the entire PUTS kit. This was necessary because pilot tests revealed that some farmers, particularly in the club good setting, refused to participate as they saw the procedure as gambling, which conflicted with their Muslim faith. Many also felt that not knowing the price created trust issues and that announcing the price would help establish trust. Since it was clear upfront that farmers' WTP would mostly fall well below the market price and given that they were bidding for a single test or just a contribution to the kit rather than the entire kit, knowing the price of the entire kit was not expected to cause anchoring, but rather to ensure high participation in the auctions.[6]

The exact procedure that we used to measure the WTP differed between the two experimental arms. In the service arm participants bid for one individualized soil test executed on their plot and included fertilizer recommendations by an expert. Prior to the bidding process, trainers provided a one-hour information session in which they introduced soil tests and explained the bidding process. Then a test run was done with a pair of house slippers to ensure that all participants understood the procedure. All participants who agreed to participate then made their bid individually and privately, one after the other. They were also asked how many tests they wanted to buy at this price if they were successful. Once all participants had made their bid and had stated the number of tests they would like to buy, a price was randomly drawn. All participants who were successful then had to make a down payment of IDR 5,000 (USD 0.32) and were given a receipt. The remaining payment, i.e., the drawn price times the number of tests, had to be paid when the soil test was actually conducted. Successful participants and the expert agreed on a date when he or she would come back to conduct the soil tests.

In the club good arm, participants bid jointly for a kit comprising 50 soil tests and took part in a group training session on how to use the kit. Public good auctions in the field are rare in the literature, with notable exceptions being Saldarriaga-Isaza et al. (2015), Carlsson et al. (2015) and Berazneva et al. (2023). As in the service arm, trainers provided a one-hour information session in each village to all farmers who attended. The enumerators explained the bidding

[5] See e.g. Berry et al. (2020) or Grimm et al. (2020) for a profound discussion of the method and its application in comparable contexts.

[6] Apart from that, farmers could also easily find out the price of the kit by googling it.

process and conducted a test run using a "blood-pressure measurement device" (another club good) to ensure that all participants had understood the procedure. Participants were allowed to discuss the bidding process before making their bids, although few did so. They then had to provide their bids individually and privately. It was also explained that the two lowest non-zero bids would be doubled.[7] Since the test kit allows up to 50 tests, each participant's bid represented a contribution to the total bid for the entire kit, not for an individual test. Participants were also asked how many soil tests they would like to perform if the group was successful. However, they could not be certain of receiving this number, as this depended on the total number of desired tests in the group, with a maximum of 50 tests available. Once all participants had made their bids, a price was randomly drawn. If the sum of all bids in that group plus the subsidy equaled or exceeded the drawn price, the group purchased the test kit at the drawn price. Each participant then paid a share of the total price (minus the subsidy) proportional to their bid relative to the total bid. For example, if a participant's bid was 10% of the total bid, they had to pay 10% of the drawn price net of the subsidy. The group had to make a down payment of IDR 5,000 (USD 0.32) times the number of participants with non-zero bids. The group and the soil test expert then agreed on a date for the delivery of the test kit and the training session. The training was scheduled for two to three hours and covered how to conduct the soil tests, including interpreting test results and their implications for fertilizer use.

4.2 Theorical considerations and hypotheses

Why should farmers attribute value to a soil test? Although the farmers in our sample had typically already been applying fertilizer for a long time, making it not a "new" technology, they tended to overuse it. Among the farmers in our sample who were also part of the SSM-project, and hence for whom we have detailed agricultural production data, we found that at least 40% overused chemical nitrogen inputs and at least 22% overused chemical phosphorus inputs. This overuse is plausibly related to the intense promotion of subsidized fertilizer and government pressure to use fertilizer during the Green Revolution (Simatupang and Timmer, 2008; Mariyono, 2014; Alta et al., 2021). Our data also suggest that prevailing social norms regarding the greenness of rice fields sometimes influence farmers to overuse chemical inputs. Farmers' farming ability is often judged by the color of their fields, with a very intense green requiring fertilizer dosages that easily exceed optimal levels (Fritz et al., 2024). Moreover, the heavy use of fertilizer in the past has contributed to deteriorating soil quality in many parts of Indonesia, including our study area, making crops more vulnerable to pests and diseases, for example, and thereby requiring adjustments in fertilizer use to maintain yields.

The farmers' problem can hence be described with the input target model (Bardhan and Udry, 1999), where the basic form of the technology is known (the use of fertilizer), but the exact target is uncertain as it depends on specific plot characteristics. A soil test generates more knowledge about the target and thus increases expected profits (Harou et al., 2017; Harou et al., 2022; Corral et al., 2020). In its most general form, the target input model reads as follows (Bardhan and Udry, 1999):

[7] From an incentive point of view, it would have been better to also account for the zero bids to encourage farmers to provide bids above zero, but we judged this as ethically not acceptable as we would have created group pressure not to make a zero bid.

If yields per hectare in season t on plot i, q_{it}, under optimal fertilizer use are normalized to 1 and k_{it} is the actual fertilizer use and κ_{it} the optimal fertilizer use (the target), the farmer's problem can be written as follows:

$$q_{it} = 1 - (k_{it} - \kappa_{it})^2, \quad (1)$$

i.e., yields are maximized if the farmer applies the optimal amount of fertilizer κ_{it}. Finding κ_{it} is a learning process in which in each season the farmer compares yields with the quantity of fertilizer used. Yet the optimal quantity is plot and time specific, i.e., there is an idiosyncratic component u_{it} since soil features and the weather, among other factors, can affect the optimal quantity that must be used on a given plot at a given time, hence:

$$\kappa_{it} = \kappa^* + u_{it}. \quad (2)$$

A soil test allows a farmer to learn about u_{it} before applying fertilizer. Hence, a soil test reduces the variance of u, σ_u^2, and thereby increases expected yields:

$$E_t(q_{it}) = 1 - \sigma_{kit}^2 - \sigma_u^2. \quad (3)$$

Soil tests are assumed to provide more accurate information about the target than extension workers who can only make general recommendations. The more heterogeneous soils are within-village communities, the higher the value from an individual soil test and the less farmers can learn from neighbors' soil tests. Perceived soil heterogeneity should therefore be a driver for the adoption of individualized soil tests (Munshi, 2004; Young, 2009; Tjernström et al., 2021; Ayalew et al., 2022; Berazneva et al., 2023).[8]

The more farmers are aware of the potential inefficiency of their fertilizer use, the more they should value soil tests. Therefore, we expected farmers who were randomly assigned to the training on sustainable soil management to reveal a higher WTP as they knew more about the importance of fertilizer management. Moreover, we anticipated farmers with higher education levels to reveal a higher WTP, as they may have better comprehended the functioning and benefits of soil tests (Foster and Rosenzweig, 1995; Ruzzante et al., 2021). Assuming age is a good proxy for experience, we also expected adoption to decrease with age, as older farmers may have perceived soil tests as less useful. Older farmers may also have found it more difficult to change their behavior.

In line with the Marshallian hypothesis (Shaban, 1987; Deininger et al., 2013), we expected farmers who owned their plots or had a fixed rent contract to reveal a higher WTP compared to those under sharecropping arrangements, as the former face a higher marginal return on investments. However, this expectation holds only if landlords cannot observe input intensities, which is not necessarily the case. Focus group discussions in our sample village communities indicated that some tenants were concerned that landlords noticed low fertilizer input, as evidenced by less green fields, and that this might jeopardize their sharecropping arrangements (Fritz et al., 2024). Finally, the WTP may decline with the number of expected soil tests other members in the farmer group undertake if farmers believe that they can use that information to

[8] Our labelling as private good (the service of a soil test) and club good (a kit of 50 tests) rests on the assumption that the information from a soil test, if shared with others, is only of limited use and that the information is not publicly shown or advertised. If the information was as beneficial for a third person as for the person owning the plot from which the soil sample was taken, and if the information was non-excludable, it would be a public good.

update their expectation about the optimal input levels on their own plots. However, this would depend on their perception of soil heterogeneity in the community.

In the club good setting the incentive structure is different. Farmers not only have the possibility to free ride on information from others, but they can also free ride on others' contributions to the soil test kit, as bids are made individually and privately. Theoretically, free riding should increase with the size of the group bidding, yet many experiments reject this hypothesis (Chaudhuri, 2011).[9] Whether free riding occurs depends on social preferences for cooperation and on participants' beliefs about others' behavior. Both lab and field studies have shown that a Nash equilibrium occurs less often than game theory would suggest. Instead, there is typically a substantial share of conditional cooperators, i.e., participants whose contributions to the public good are positively correlated with their beliefs about their group members' contributions (Fischbacher et al., 2001; Chaudhuri, 2011; Vesterlund, 2017). For example, Fischbacher et al. (2001) found that 50% of their participants in a public good game were conditional cooperators. Some studies have also shown that contributions in public good games are driven by distributional concerns, i.e., participants contribute more if they are inequity averse (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). In our case, the group setting as opposed to the individual setting allowed for the distribution of costs according to income, either explicitly or implicitly, which should address affordability issues at the lower end of the distribution. We expected this to occur if farmers shared some egalitarian norms.

Moreover, the literature shows that cooperation in public good games is enhanced if participants communicate beforehand (see e.g., Isaac and Walker, 1988; Cason and Khan, 1999; Bochet et al., 2006; Brosig et al., 2006; Eisenkopf, 2018). We did not prevent participants from discussing their contributions before making their bids individually and privately. Hence, we expected groups utilizing this option to make, on average, higher bids.[10] Given that the option for "communication" was not randomized, we could not control whether groups chose to communicate in the first place and thus cannot make causal claims regarding its impact on bid size.

In both experimental settings, the private good and the club good setting, participants were asked to indicate the number of tests they wished to conduct. In the private good setting, participants simply stated how many tests they would like to purchase at the drawn price if their bid was successful. In the club good setting, participants had to state how many tests they expected to perform with their contribution if the group was successful. Given that participants in the club good setting a priori did not know how many tests other group members expected to perform, there was a substantial amount of uncertainty involved. Yet we expected participants who wanted to conduct more tests to make higher contributions. In both settings, we anticipated that for a given valuation of the information, the number of desired tests would increase with the number of plots participants cultivate. Participants may also attribute a higher value to the first test than to subsequent ones, in which case the per test bid should decrease as the number of desired tests increases.

[9] In the experiment in Malawi, the authors found that farmers exhibited cooperative behavior, but as group size increased by five (the median group size), average individual contributions dropped by 7% and the coefficient of variation of contributions within the group increased by 41% (Berazneva et al., 2023). This was also confirmed by an earlier study in the same context (Nourani et al., 2021).

[10] In the private good setting, participants could also discuss the auction while waiting their turn, but there it should matter less for their bids.

Public good experiments like ours in the field are rare, as most studies rely on lab experiments, which typically imply a high internal validity but might be weak regarding their external validity. Notable exceptions include Carlsson et al. (2015), who explored the willingness to contribute to a bridge in a rural community in Vietnam, and Saldarriaga-Isaza et al. (2015), who studied a collective action to ban mercury amalgamation in artisanal gold mining in Colombia. Hence, we believe our study makes an innovative contribution to that literature as well.

5. Results

5.1 A description of the outcome of the BDM auction

We begin by focusing on the revealed WTP for the individual soil test offered as a service (Table 2). The average WTP was IDR 15,600 (USD 0.99), which is a considerable amount and corresponds to 43% of the cost of a single soil test (excluding the costs of performing the test). The highest bid amounted to IDR 57,900, which is 1.6 times the cost of the test. Among all respondent in the service arm, 25.4% did not make a positive bid (zero bid). Those who made a positive bid stated that they wanted to purchase an average of 1.7 tests at their bid price.

In the club good arm, the average WTP was IDR 24,300 (USD 1.54), which is higher than in the service arm. However, this bid represents the total contribution a participant makes towards the cost of the entire kit of 50 tests. Therefore, to obtain a per test WTP, this contribution had to be divided by the number of tests a participant expected to obtain. Participants could of course not be certain of obtaining their desired number of tests, as this depended on how many tests other group members wanted to perform. On average, participants expected to be able to perform 2.1 tests, reducing the per test WTP to IDR 14,500, which is slightly less than the average WTP in the service arm. However, the share of zero bids was 22% higher in the club good arm compared to the service arm (31.1% vs 25.4%), possibly suggesting some free riding on others' bids in this setting. This will be explored further below.

The average group size, i.e., the number of farmers in each village participating in the bidding process, was 14.7 in the club good arm and 12.3 in the service arm. While the absolute spread of the bids was comparable between the two settings, the club good arm exhibited a higher standard deviation.

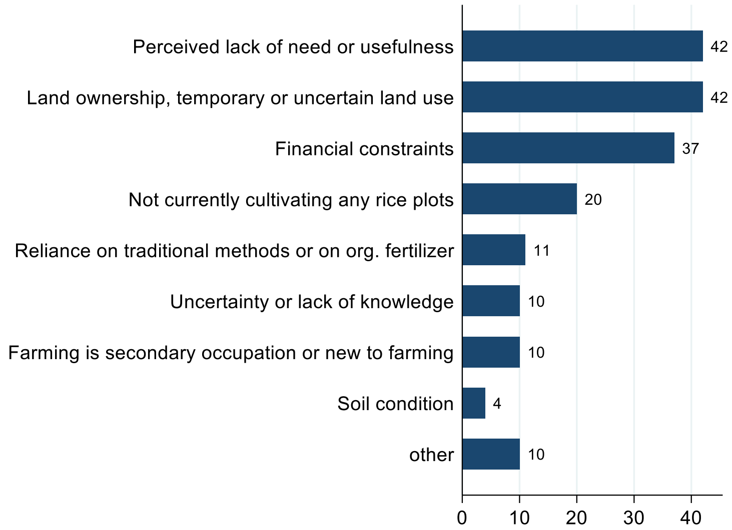
When we directly asked participants with zero bids about their reasons, weak beliefs in the technology, land rental arrangements, and affordability emerged as dominant factors (see Figure 4). There were no substantial differences in the reasons given between the two experimental arms.

Table 2: Outcome of BDM auction

	Service	Club Good
Individual level		
WTP (total contribution) (in 1,000 IDR)	15.6	24.3
	(21.0)	(31.3)
Desired # of tests per participant (if non-zero bid)	1.7	2.1
	(1.1)	(1.4)
WTP per soil test (in 1,000 IDR)	15.6	14.5
	(21.0)	(21.1)
Village (Dusun) level		
Group size	12.3	14.7
	(3.4)	(3.9)
Spread of WTP per test within group (in 1,000 IDR)	56.5	57.2
	(40.1)	(50.6)
Group min WTP per test (in 1,000 IDR)	1.5	0.9
	(4.0)	(2.4)
Group max WTP per test (in 1,000 IDR)	57.9	58.1
	(39.4)	(50.3)
Share of zero bids by group	0.254	0.311
	(0.227)	(0.267)
Share of partic. discuss. bid with others bef. bidding	0.271	0.338
	(0.445)	(0.474)
Outcome of BDM auction		
Successful (share individuals / groups)	0.525	0.619
Average bid if successful (in 1,000 IDR)	27.0	437.5
	(23.7)	(111.6)
Participants	292	305
Village groups (Dusuns)	24	21

Notes: Standard deviations in parentheses.

Figure 4: Self-reported reasons for zero bids (N=186)



Surprisingly, in the club good setting, only 33.8% of participants discussed their bids with others before submitting them (see Table 2). As expected, in the service setting, this share was lower, but not much (27.1%). In most cases, once all instructions had been given, participants simply remained seated and waited for their turn to make a bid.

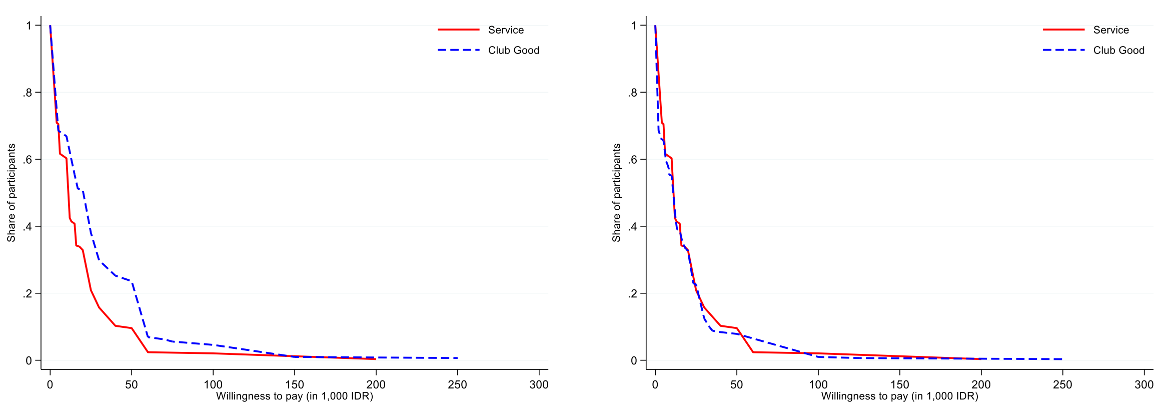
Since we drew prices well below the market price, 52.5% of all bids in the service good arm were successful, while 61.9% of all groups in the club good arm were successful (see Table 2).

5.2 Demand curves

Figure 5 presents demand curves showing the uptake of soil tests as a function of the price. The left figure uses gross price bids, while the right figure uses the adjusted per test price bids. The gross price uptake was higher in the club good arm for any price between IDR 15,000 and IDR 150,000. However, this was not the case with adjusted prices, where the curves largely overlap. At the actual cost of IDR 36,000, the per test uptake would be 20%.

Figure 5: Demand curves

(a) gross price bids (b) adjusted price bids (club good)



5.3 *Regression-based analysis of the WTP*

Guided by our theoretical considerations, we now analyse the determinants of the WTP in both experimental arms in more detail using regressions. We specifically examine whether prior training affected farmer's valuation of soil tests. To do this, we estimate a simple linear regression of the following form and pool the observations from both experimental arms:

$$WTP_i = \beta_0 + \beta_1 ClubGood_i + \beta_2 SoilTraining_i + \beta_3 NewInSample_i + X_i'\gamma + u_i, \qquad (4)$$

where the variable *ClubGood* takes the value one if participant i was in the club good experimental arm and zero if the participant was in the service good arm. We introduce a variable indicating exposure to the one-day training in the SSM-project and controll for the fact that some participants were not covered by the SSM-project and, therefore, had neither been trained nor surveyed. Being surveyed, even without training exposure, may have increased awareness of soil issues, potentially affecting the revealed WTP. Furthermore, vector X_i' includes education, age, group size, and whether participants discussed their bid with others before being asked about it. In a separate regression we also test whether the type of tenancy influences WTP. Furthermore, we run a regression in which we replace the dependent variable with the number of desired tests. We cluster standard errors at the village (dusun) level to account for possible within-village correlation of observations. Tables 3 and 4 show the results.

In Table 3, Col. (1) we use gross bids and in Cols. (2) to (6) adjusted price bids. The latter assumes that farmers correctly anticipated how many tests they would be able to conduct.[11] Col. (1) shows that when using the gross bid, bids are higher in the club good arm. The difference is statistically significant, amounting to IDR 8,441 or 54.1% (8,441/15,600). However, when we estimate the model with the adjusted bid in the club good arm (Col. (2)), we find that this

[11] If they anticipated that the total number of desired tests within their group would exceed the maximum number of 50, and hence that they would ultimately only be able to perform fewer tests than indicated, the difference would lie between the estimates in Col. (1) and Col. (2). In our sample, the average total number of tests within a group is 20.6, with a maximum of 40 and a minimum of 5, but this information was, of course, not known to the farmers when making their bid.

difference shrinks to IDR 785 and is no longer statistically significant, indicating that the average bid per soil test does not differ between the two settings.

Table 3. Treatment effects and correlates of the WTP for soil testing

	(1) WTP Gross	(2) WTP Adj. CG	(3) WTP Adj. CG	(4) WTP Adj. CG	(5) WTP Adj. CG	(6) WTP Adj. CG
Club good (=1)	8.441***	0.785	-9.715	-8.696	-9.856	-6.496
	(2.921)	(2.038)	(7.279)	(7.498)	(7.250)	(7.509)
No prior train., not surveyed (=1)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
No prior training, surveyed (=1)	8.397**	1.899	2.109	2.096	2.137	1.984
	(3.317)	(2.441)	(2.443)	(2.382)	(2.346)	(2.377)
Prior training, surveyed (=1)	8.760***	0.879	0.844	0.825	0.786	0.264
	(2.894)	(2.160)	(2.107)	(2.073)	(2.035)	(2.026)
Group size		-0.839***	-1.301***	-1.285***	-1.301***	-1.124***
		(0.224)	(0.329)	(0.337)	(0.329)	(0.319)
Group size x Club good			0.745*	0.660	0.634	0.524
			(0.442)	(0.457)	(0.460)	(0.454)
Age	0.234	0.110	0.103	0.109	0.109	0.105
	(0.154)	(0.120)	(0.122)	(0.121)	(0.120)	(0.117)
Primary or lower (=1)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Junior High (=1)	9.970***	5.654**	5.525**	5.548**	5.434**	4.391*
	(3.636)	(2.276)	(2.254)	(2.245)	(2.182)	(2.211)
Sen. High or higher (=1)	12.359***	8.492***	8.122***	7.812***	7.758***	6.619***
	(2.740)	(2.399)	(2.377)	(2.338)	(2.324)	(2.380)
Communication beforehand (=1)				4.149**	1.316	0.697
				(1.974)	(1.812)	(1.889)
Comm. bef. x Club good					5.276	6.064
					(3.630)	(3.639)
Desired # tests						3.173***
						(1.011)
Desired # tests x Club good						-2.059
						(1.235)
Surv. vs. training (*p*-value)	0.915	0.616	0.528	0.524	0.500	0.405
R2	0.072	0.055	0.059	0.067	0.070	0.088
Obs.	596	596	596	596	596	596

Notes: Standard errors (in parentheses) clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. The outcome (WTP) in cols. (1) to (6) is measured in IDR 1,000.

The exposure to prior training does not affect the price bids. Although we see an effect on the unadjusted bids in comparison to new participants in Col. (1), this effect is not different between farmers who were exposed to training and those not exposed to training under the SSM-project. Furthermore, adjusting the bids renders even this effect insignificant.

In line with our hypothesis, we observe that the WTP increases with formal education. Participants with junior secondary schooling compared to those with no and primary schooling, made bids that were higher on average by about IDR 5,654 (Col. (2)). For participants with senior high schooling, bids were even higher by IDR 8,492, which is more than 50% of the mean. This may also capture a wealth effect that we cannot control for. We do not find any correlation with age, conditional on all other controls, indicating that older participants neither value soil tests more nor less.

Group size matters and is consistently associated with a negative effect. The regression in Col. (2) suggests that with each additional member, the bid decreases by IDR 839. As outlined in our theoretical considerations, a possible explanation could be that participants anticipated that in a larger group, more farmers around them would perform a soil test from which they could also learn, thus reducing the value they attached to a private soil test. In the club good arm, the negative coefficient could also imply that participants free ride on others' contributions. However, if we re-estimate the model with an interaction term between the experimental arm and group size (Col. (3)), we find that the effect is actually less than half the size in the club good arm compared to the service arm (-1.301 vs. -1.30 + 0.745 = -0.556), suggesting that free riding on others' contributions in the club good arm is not the key driver of the group size effect. This could be due to generally high trust levels within farmer groups.

In Col. (4), we add an indicator variable for whether participants discussed their bid with others beforehand. We find that participants who did discuss beforehand made bids that were IDR 4,149 higher. This is a very sizable effect and aligns with outcomes observed in many public good games performed in the lab (see Section 4). Indeed, Col. (5) suggests that this effect is almost entirely driven by participants in the club good setting (who also engaged a bit more in communication than those in the private good setting, 33.8% vs. 27.1%), yet the interaction effects are not very precisely estimated. A joint F-test of the linear and interaction effect yields a p-value of 0.102.

In Col. (6), we add the number of desired tests as a control for both experimental arms. We find that for participants in the service arm the bid per test increases by IDR 3,173 with each additional desired test, suggesting that farmers who want to carry out more tests also have a higher WTP. For participants in the club good arm this effect seems to be much lower. The interaction effect is not very precisely estimated (p-value: 0.103) but the direction and size of the coefficient hint to a smaller effect for this treatment arm. This may indicate that farmers in the club good arm, attribute a lower value to additional tests, maybe because they anticipate that they can also learn from tests other farmers in their group have carried out.

Since, we re-assigned five villages from the club good treatment to the service treatment because in these fiver villages less than ten farmers attended the experiment, we re-estimated all regressions shown in Table 3 without the re-assigned villages. This reduces the sample by 49 observations. The inclusion of these villages could have biased our results given that their treatment status is not entirely exogenous. The results are shown in Table B.1. The results are qualitatively the same. The only difference is that the group size effect is not anymore significant if adjusted bids are used., yet it is still negative throughout, but about 50% smaller.

In Table 4 we use the number of desired tests rather than the price bid as the dependent variable. We run two regressions, one with participants with zero bids (col. (1)) and one excluding them (col. (2)). We observe that non-zero bidding participants in the club good experimental arm requested, on average, 1.5 more tests than those in the private good experimental arm (col. (2)). This effect is smaller and not significant if all participants are considered. This is largely due to the fact that there are more zero bids in the club good setting. We also observe that exposure to the soil management training increases the number of desired tests by 0.31 to 0.42 tests. Because these estimates lack precision, they are not statistically different from the pure survey effect (*p*-value between 0.34 and 0.50) the difference in the effect size strongly suggest that the training had an additional effect, increasing the number of tests participants want to carry out.

Table 4. Treatment effects and correlates of the WTP for soil testing

	(1) Desired # of tests	(2) Desired # of tests
Club good (=1)	0.514	1.583***
	(0.387)	(0.008)
No prior training, not surveyed (=1)	Ref.	Ref.
No prior training, surveyed (=1)	0.242	0.202
	(0.106)	(0.181)
Prior training, surveyed (=1)	0.419**	0.312*
	(0.032)	(0.075)
Group size	-0.046	0.009
	(0.158)	(0.743)
Group size x Club good	-0.014	-0.083**
	(0.738)	(0.031)
Age	0.004	-0.001
	(0.421)	(0.842)
Primary or lower (=1)	Ref.	Ref.
Junior High (=1)	0.438**	0.007
	(0.024)	(0.974)
Sen. High or more (=1)	0.589***	0.224
	(0.001)	(0.231)
Surv. vs. training (*p*-value)	0.350	0.516
R2	0.078	0.068
Obs.	596	416

Notes: In col. (2) respondents who made zero bids are excluded from the estimation.

We also explore whether the WTP varies with tenancy status. The Marshallian inefficiency hypothesis predicts that owned plots and plots under fixed rent contracts receive higher input intensities than sharecropped plots due to higher marginal returns, especially if landowners cannot monitor inputs. This should also apply to soil tests. We can test this hypothesis in the sub-sample of farmers who made non-zero bids, as we have tenancy status information on each plot they wanted to test. Our results suggest that WTP indeed decreases with an increase in the share of land under sharecropping that the farmer wants to test. Specifically, a ten-percentage point increase in the share of land under sharecropping reduces the adjusted WTP by IDR 407. These results are shown in the appendix (Table C.1). However, farmers who own or rent under fixed rent contracts may differ from those who rent under sharecropping arrangements. Therefore, we re-estimate the model for a sub-sample of farmers who do both. This information is available for all farmers who were also surveyed in the SSM-project. The coefficient for the WTP is again negative, but not statistically significant, likely due to the very limited power of this small sub-sample. Overall, the results suggest that sharecropping reduces the WTP for soil testing, but this effect is not very robust and should not be interpreted in a causal way, it is suggestive evidence at best.

6. A comparison of soil test provision modes and the role of subsidies

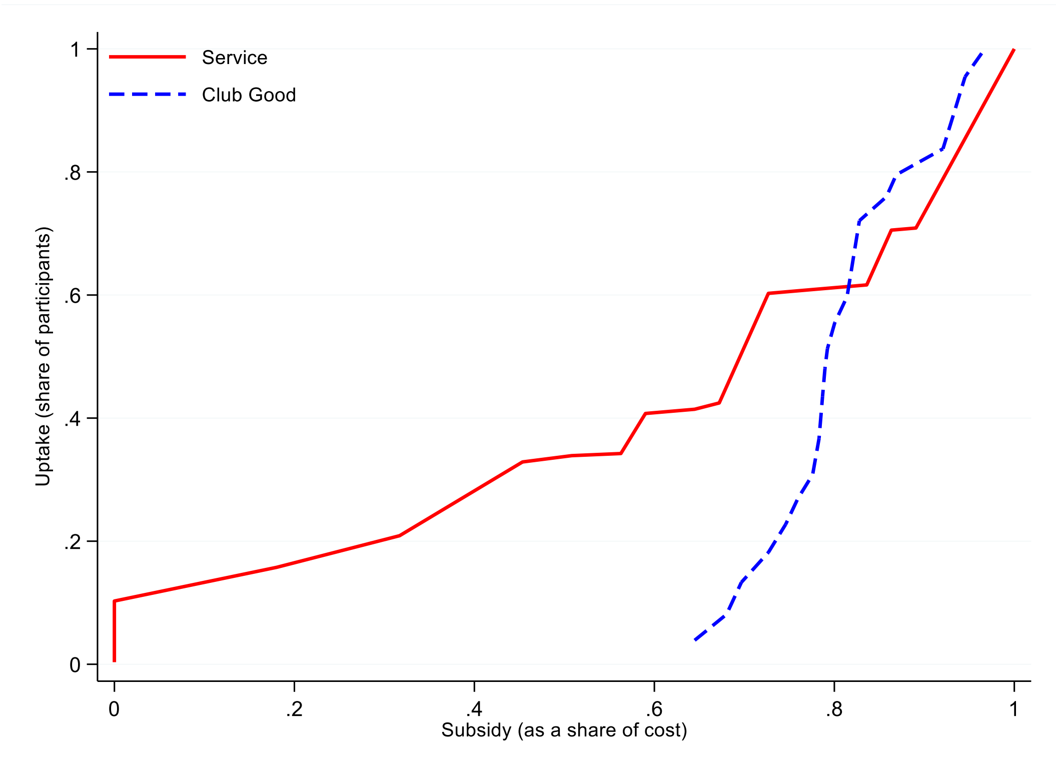
We did not expect the revealed price bids to be cost covering; rather, we aimed to obtain a robust idea of the extent to which farmers could contribute to the total costs, enabling conventional extension services to offer soil tests at a subsidized price. Figure 6 shows uptake as a function of the share of total costs subsidized for both experimental settings. We account only for the direct costs of the test and do not include the cost for the person who delivers the service or the training assuming that this could be integrated into the routines of extension workers without additional costs. For lower subsidies, providing soil tests as an individual service is the most effective dissemination mode. Group uptake is zero unless at least 60% of the costs are subsidized. However, for subsidies above 75%, i.e., a farmer's contribution is 25% or less, uptake is higher in the group setting. With an 80% subsidy, uptake is about 60% in the service arm and 70% in the club good arm. At a 90% subsidy, the difference in uptake increases to almost 20 percentage points.

Subsidizing soil tests can be justified on several grounds. First, reducing chemical fertilizer use can prevent soil degradation, decrease plant vulnerability to pests and diseases, and reduce ground water pollution (Stevens et al., 1999; Bijay-Singh and Craswell, 2021). Hence, subsidies can help internalize these externalities. Second, the information from a soil test, if shared, also carries valuable information for others, even if that value decreases with soil heterogeneity within farmer groups.

Given the potentially high private returns of these low-cost soil tests, it may also be possible to disseminate them at cost-covering prices once farmers have learned how to use them, how to follow the recommendations, and have observed their profitability. ISRIS's demonstration project in which plots were alternatively cultivated as usual or according to the soil test recommendation, suggested that yields could increase by 6% to 13% if the recommendations are followed (Irawan et al., 2021). Since this soil test provides valid recommendations for at least three seasons, investing in a soil test at its full price is likely to be cost effective.

Soil tests have been shown to be potentially cost effective in other settings as well (see e.g., Murphy et al., 2020). If farmers gradually realize that soil tests are profitable, it should be possible to reduce subsidies without lowering uptake. Thus, a subsidized start may come with self-sustaining use later.

Figure 6: Uptake as a function of the subsidized costs



Notes: Costs include only the costs of the soil tests without the service and training delivered by an extension agent, i.e., 36,600 IDR in the service arm and 1,830,000 IDR in the club good arm. In the club good arm, uptake is weighted by group size, i.e., both lines show uptake at the individual level.

Soil tests have been shown to be potentially cost effective in other settings as well (see e.g., Murphy et al., 2020). If farmers gradually realize that soil tests are profitable, it should be possible to reduce subsidies without lowering uptake. Thus, a subsidized start may come with self-sustaining use later.

7. Conclusion

Our analysis shows that farmers attribute substantial value to information on optimal fertilizer use. Farmers are willing to pay an average of USD 1 for a soil test, which corresponds to 43% of its actual cost. Among those who made a positive bid, farmers wanted to test 1.7 plots, which is, on average, more than half of the plots they cultivate. Those who made zero bids reported that the main reasons were their belief that soil tests are not useful, land ownership issues, and financial constraints. Although we do not find an effect of exposure to the one-day training on sustainable land management practices on the WTP, we find some suggestive evidence that it increased the number of tests participants wanted to carry out. Higher education was associated with a higher WTP, which may suggest that education helps in understanding the benefits of soil tests; however, it could also partially reflect a wealth effect. We found suggestive evidence that tenure status matters too. Sharecropped plots imply a lower WTP for soil tests, but this effect is not very robust and should not be interpreted in a causal way. Moreover, the qualitative field work suggested that in the sampled village communities landlords can often monitor the effort level of their tenants.

We found an almost identical WTP on a per test basis when we auctioned soil test kits of 50 tests to farmer groups, compared to offering serviced tests to individual farmers. However, non-zero bidding participants in the club good setting demanded a slightly higher number of tests. Consistent with evidence from many public good games conducted in the lab (see e.g., Fischbacher et al., 2001; Chaudhuri, 2011; Vesterlund, 2017), we found little evidence of free riding in our field experiment, suggesting that a Nash equilibrium is not the norm. We also found that communication among participants before the bidding is associated with higher bids, aligning with evidence from many public good games conducted in the lab (see e.g., Isaac and Walker, 1988; Cason and Khan, 1999; Bochet et al., 2006; Brosig et al., 2006; Eisenkopf, 2018). Qualitative interviews conducted in the targeted communities after the experiment suggested that the group setting in the club good arm may have fostered a sense of joint responsibility and mutual support, possibly counteracting incentives to free ride on others' contributions.

Our experiment also demonstrated that integrating soil tests with existing extension services could be relatively straightforward. Subsidies can be justified by the potential environmental benefits of preventing fertilizer overuse and improving soil management, yet they could be removed gradually as private returns should compensate for the costs. Balew et al. (2024) show convincingly that full subsidies can indeed enhance subsequent longer-term adoption and WTP for integrated pest management.

In our setting, at lower subsidies, providing soil tests as an individual service is most effective, whereas at higher subsidies, offering entire test kits and training is more effective in enhancing uptake. The latter might also be more sustainable long-term, as it may increase the likelihood that farmer groups integrate soil testing into their group activities. It also allows poorer farmers to benefit from such tests, since contributions can be made according to their ability to pay. A downside of the group setting is that some farmers seem reluctant to approach the person who is storing the kit, typically the farmer group head. A follow-up survey five months after the experiment revealed that indeed, among farmers who participated in the service experimental arm, the share of soil tests already completed was higher.

Future research should focus on the impact of low-cost soil tests on actual input use, yields, and long-term soil health. So far, there is little evidence regarding these tests' environmental impact and whether they can help to achieve the twin goal of higher profits and reduced environmental degradation.

Appendix A: Sampling procedure in the SSM-project

To sample villages and households in the SSM-project, we applied a multi-stage sampling design to sample 1,104 respondents from 69 villages, i.e., 16 respondents per village. In the first stage, we dropped unsuitable villages from the village database for the three districts - Sleman, Kulon Progo and Bantul. In the second stage, we randomly sampled one sub-village from each of the remaining villages (sub-village, or *Dusun* in Bahasa, is an administrative subdivision of a village in Indonesia). A team of enumerators then visited each sub-village to obtain more detailed information on farmers' cultivation focus and demography from the sub-village head and the farmer group head. Two-hour information sessions about digital agricultural resources were then provided in all the sampled sub-villages.

In the third stage, we randomly sampled respondents among the attendees of the information session. The three sampling criteria for respondents were as follows: (1) access to a smartphone, (2) rice cultivation of at least one plot, and (3) younger than 65 years (70 years in the case of the farmer group head). In some information sessions there were fewer than 16 suitable respondents. In such cases, we asked the farmer group head to provide us with contact details of other farmers who had not attended the information session.

Access to a smartphone was important in the SSM-project as the soil management training worked with offline and online resources (platform services) to provide information.

Appendix B: Robustness check

Table B.1. Treatment effects and correlates of the WTP for soil testing, excluding reassigned villages

	(1) WTP Gross	(2) WTP Adj. CG	(3) WTP Adj. CG	(4) WTP Adj. CG	(5) WTP Adj. CG	(6) WTP Adj. CG
Club good (=1)	10.030***	1.584	-1.535	1.134	-0.209	4.235
	(2.885)	(2.022)	(9.146)	(9.482)	(8.896)	(9.023)
No prior training, not surveyed (=1)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
No prior training, surveyed (=1)	8.662**	2.618	2.678	2.599	2.651	2.625
	(3.408)	(2.460)	(2.443)	(2.346)	(2.333)	(2.381)
Prior training, surveyed (=1)	7.748**	0.422	0.495	0.379	0.387	-0.186
	(2.945)	(2.048)	(2.071)	(2.020)	(2.000)	(1.951)
Group size		-0.601**	-0.758	-0.652	-0.697	-0.489
		(0.237)	(0.455)	(0.468)	(0.446)	(0.438)
Group size x Club good			0.214	0.018	0.040	-0.097
			(0.550)	(0.570)	(0.559)	(0.553)
Age	0.266	0.152	0.150	0.159	0.158	0.146
	(0.159)	(0.123)	(0.125)	(0.124)	(0.123)	(0.120)
Primary or lower (=1)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Junior High (=1)	11.000***	6.604***	6.559***	6.673***	6.558***	5.674**
	(3.879)	(2.391)	(2.391)	(2.387)	(2.322)	(2.320)
Sen. High or higher (=1)	12.590***	8.664***	8.541***	8.237***	8.165***	6.626**
	(2.883)	(2.515)	(2.539)	(2.507)	(2.487)	(2.568)
Communication beforehand (=1)				5.200**	3.203*	2.182
				(2.036)	(1.737)	(1.785)
Comm. bef. x Club good					3.458	4.686
					(3.505)	(3.486)
Deisred # tests						3.946***
						(1.143)
Desired # tests x Club good						-2.801**
						(1.307)
Surv. vs. training (p-value)	0.801	0.299	0.783	0.290	0.285	0.197
R2	0.083	0.046	0.046	0.059	0.061	0.083
Obs.	547	547	547	547	547	547

Appendix C: Effect of tenancy status on the WTP for soil testing

Table C.1: Effect of tenancy status on the WTP for soil testing

	(1)	(2)	(3)
	Only participants with WTP>0		Only participants with WTP>0 and who have own/fixed rent and sharecropping
	WTP Gross	WTP Adj. CG	WTP Adj. CG
Share of land to be tested	-4.567*	-4.066*	-2.429
under sharecropping	(2.569)	(2.274)	(6.354)
(vs. owned and fixed rent)			
Controls	Yes	Yes	Yes
R2	0.098	0.039	0.072
Obs.	416	416	111

Notes: Standard errors (in parentheses) clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1. Controls used: experimental arm, group size, age, education and status in SSM-project.

References

Alta, A., I. Setiawan and A.N. Fauzi (2021), Beyond Fertilizer and Seed Subsidies: Rethinking Support to Incentivize Productivity and Drive Competition in Agricultural Input Markets, Policy Paper, No. 43, Center for Indonesian Policy Studies (CIPS), Jakarta.

Ayalew, H., J. Chamberlin and C. Newman (2022), Site-specific agronomic information and technology adoption: A field experiment from Ethiopia. *Journal of Development Economics*, 156, 102788.

Balew, S., E. Bulte and K. Menale (2024), Short-Run Subsidies and Long-Run Willingness to Pay: Learning and Anchoring in an Agricultural Experiment in Ethiopia. *American Journal of Agricultural Economics*, 1–15.

Bardhan, P. and C. Udry (1999), *Development Microeconomics*. Oxford University Press: Oxford.

Beaman, L., D. Karlan, B. Thuysbaert and C. Udry (2013), Profitability of fertilizer: Experimental evidence from female rice farmers in Mali. *American Economic Review*, 103 (3): 381-386.

Beg, S., M. Islam and K.W. Rahman (2024), Information and behavior: Evidence from fertilizer quantity recommendations in Bangladesh. *Journal of Development Economics*, 166, 103195.

Berazneva, J., A. Maertens, W. Mhango and H. Michelson (2023), Paying for agricultural information in Malawi: The role of soil heterogeneity. *Journal of Development Economics*, 165, 103144.

Berry, J., G. Fischer and R.P. Guiteras (2020), Eliciting and utilizing willingness to pay: Evidence from field trials in Northern Ghana. *Journal of Political Economy*, 128 (4): 1436- 1473.

Bijay-Singh and E. Craswell (2021), Fertilizers and nitrate pollution of surface and ground water: an increasingly pervasive global problem. *SN Applied Sciences*, *3* (4): 518.

Bochet, O., T. Page and L. Putterman (2006), Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior and Organization*, 60 (1): 11–26.

Bold, T., K.C. Kaizzi, J. Svensson and D. Yanagizawa-Drott (2017), Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in Uganda. *Quarterly Journal of Economics*, 132 (3): 1055-1100.

Bolton, G. and A. Ockenfels (2000). ERC—a theory of equity, reciprocity and competition. *American Economic Review*, 90, 166–193.

Brosig, J., J. Weimann and A. Ockenfels (2003), The Effect of Communication Media on Cooperation. *German Economic Review*, 4 (2): 217–241.

Carlsson, F., O. Johansson-Stenman and P. Khanh Nam (2015), Funding a new bridge in rural Vietnam: A field experiment on social influence and default contributions. *Oxford Economic Papers*, 67 (4): 987-1014.

Cason, T.N. and F.U. Khan (1999), A laboratory study of voluntary public goods provision with imperfect monitoring and communication. *Journal of Development Economics*, 58 (2): 533–552.

Chaudhuri, A. (2011), Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14, 47–83.

Conley, T. and C. Udry (2001), Social learning through networks. The adoption of new agricultural technologies in Ghana. *American Journal of Agricultural Economics*, 83 (3): 668-673.

Conley, T. and C. Udry (2010), Learning about a new technology: Pineapple in Ghana. *American Economic Review*, 100 (1): 35-69.

Corral, C., X. Giné, A. Mahajan, E. Seira (2020), Autonomy and Specificity in Agricultural Technology Adoption: Evidence from Mexico. NBER working paper #w27681), National Bureau of Economic Research, Boston, MA.

Deininger, K., J. Songqing and V. Yadav (2013), Does Sharecropping Affect Long-term Investment? Evidence from West Bengal's Tenancy Reforms. *American Journal of Agricultural Economics, 95 (3): 772-779.*

Duflo, E., M. Kremer and J. Robinson (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, 101 (6): 2350-2390.

Eisenkopf, G. (2018), The long-run effects of communication as a conflict resolution mechanism. *Journal of Economic Behavior and Organization*, 154: 121–136.

Fehr, E. and K. Schmidt (1999), A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114, 817–868.

Fischbacher, U., S. Gächter and E. Fehr (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3), 397–404.

Fishman, R, J. Gars, A. Kishore, Y. Rothler and P. Ward (2019), Confidence and Information Usage: Evidence from Soil Testing in India. Working Paper (https://jaredgars.github.io/papers/Submission_confidence.pdf).

Foster, A. and M. Rosenzweig (1995), Learning by Doing and Learning form Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy*, 104: 1176-1209.

Fritz, M., N. Luck, N.H. Primasari and U. Sawhney (2024), My opinion, your opinion - Do group perceptions prevent the adoption of sustainable farming practices? Working Paper.

Grimm, M., L. Lenz, J. Peters and M. Sievert (2020), Demand for off-grid solar electricity – Experimental evidence from Rwanda. *Journal of the Association of Environmental and Resource Economists*. 7 (3): 417-454.

Grimm, M. and N. Luck (2023), Experimenting with a green 'Green Revolution'. Evidence from a randomised controlled trial in Indonesia. *Ecological Economics,* 205, 107727.

Harou, A.P., Y. Liu, C. Barrett and L. You (2017), Variable returns to fertiliser use and the geography of poverty: Experimental and simulation evidence from Malawi. *Journal of African Economies*, 26 (3): 342–371.

Harou, A.P., M. Madajewicz, H. Michelson, C.A. Palm, N. Amuri, C. Magomba, J.M. Semoka, K. Tschirhart and R. Weil (2022), The joint effects of information and financing constraints on technology adoption: Evidence from a field experiment in rural Tanzania. *Journal of Development Economics*, 155, 102707.

Hazell, P.B.R. (2009), The Asian green revolution. IFPRI Discussion Paper 00911, IFPRI, Washington D.C.

IFAD (2013), *Smallholders, Food Security, and the Environment*. International Fund for Agricultural Development (IFAD), Rome.

Isaac, R.M., J.M. Walker (1988), Communication and Free-riding Behavior: The Voluntary Contribution Mechanism. *Economic Inquiry*, 26 (4): 585–608.

Islam, M. and S. Beg (2021), Rule-of-Thumb Instructions to Improve Fertilizer Management: Experimental Evidence from Bangladesh. *Economic Development and Cultural Change*, 70 (1): 237-281.

Jack, B.K. (2011), Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries. CID Research Fellow and Graduate Student Working Paper Series 2011.50, Harvard University, Cambridge.

Jing, L. and R. Cheo (2013), House money effects, risk preferences and the public good game. *Economic Letters*, 120 (2), 310-313.

Kokoye, S.E.H., C.M. Jolly, J.J. Molnar, D.A. Shannon and G. Huluka (2018), Farmer Willingness to Pay for Soil Testing Services in Northern Haiti. *Journal of Agricultural and Applied Economics*, 50 (3): 429–451.

Krishnan, P. and M. Patnam (2014), Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption? *American Journal of Agricultural Economics*, 96(1), 308-327.

Lai, W. (2017), Pesticide use and health outcomes: evidence from agricultural water pollution in China. *Journal of Environmental Economics and Management*. 86, 93–120.

Langyintuo, A. (2020), Smallholder Farmers' Access to Inputs and Finance in Africa. In: Gomez y Paloma, S., Riesgo, L., Louhichi, K. (eds.), *The Role of Smallholder Farms in Food and Nutrition Security*. Springer, Cham.

Lee, D.R. (2005), Agricultural Sustainability and Technology Adoption: Issues and Policies for Developing Countries. *American Journal of Agricultural Economics*, 87: 1325-1334.

Magruder, J.R. (2018), An assessment of experimental evidence on agricultural technology adoption in developing countries. *Annual Review Resource Economics*, 10 (1), 299–316.

Mariyono, J. (2014), Rice production in Indonesia: policy and performance. *Asia Pacific Journal of Public Administration*, 36 (2): 123-134.

Michler, J.D., K. Baylis, M. Arends-Kuenning, K. Mazvimavi (2019), Conservation agriculture and climate resilience. *Journal of Environmental Economics and Management*, 93: 148–169.

Munshi, K. (2004), Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, 73 (1), 185–213.

Murphy, D.M.A., D. Roobroeck, D.R. Lee and J. Thies (2020), Underground Knowledge: Estimating the Impacts of Soil Information Transfers Through Experimental Auctions. *American Journal of Agricultural Economics*, 102: 1468-1493.

Ruzzante, S., R. Labarta and A. Bilton (2021), Adoption of agricultural technology in the developing world: A meta-analysis of the empirical literature. *World Development*, 146, 105599.

Saldarriaga-Isaza, A., C. Villegas-Palacio and S. Arango (2015), Phasing out mercury through collective action in artisanal gold mining: Evidence from a framed field experiment. *Ecological Economics*, 120: 406-415.

Shaban, R.A. (1987), Testing between Competing Models of Sharecropping. *Journal of Political Economy*, 95 (5):893-920.

Simatupang, P. and P.C. Timmer (2008), Indonesian rice production: policies and realities. *Bulletin of Indonesian Economic Studies*, *44*(1): 65-80.

Sukayat, Y., I. Setiawan, U. Suharfaputra and G. Kurnia (2023), Determining Factors for Farmers to Engage in Sustainable Agricultural Practices: A Case from Indonesia. *Sustainability*, 15, 10548.

Tamim, A., A.P. Harou, C. Magombab, H. Michelson and C. Palm (2020), The Long-Term Effects of Relaxing Information and Credit Constraints on Adoption, Retention, and Soil Perceptions: Evidence from a Randomized Experiment in Tanzania. 2020 Annual Meeting, July 26-28, Kansas City, Missouri 304604, Agricultural and Applied Economics Association.

Tjernström, E., T.J. Lybbert, R.F. Hernández and J.S. Correa (2021), Learning by (virtually) doing: Experimentation and belief updating in smallholder agriculture. *Journal of Economic Behavior and Organization*, 189, 28–50.

Vesterlund, L. (2017), Voluntary Giving to Public Goods: Moving Beyond the Linear VCM. In: Kagel, John H., Roth, Alvin E. (Eds.), *The Handbook of Experimental Economics*, Vol. 2. Princeton University Press, Princeton, NJ.

Widowati, L.R., D. Nursyamsi, S. Rochayati and M. Sarwani (2011), Nitrogen management on agricultural land in Indonesia. In International Seminar on Increased Agricultural Nitrogen Circulation in Asia: Technological Challenge to Mitigate Agricultural N Emissions (pp. 181-195).

Widowati, L.R., D. Setyorini, A. Kasno, Nurjaya, SM. Nanan, S. Supardi and Sulaeman (2004), Pengembangan soil test kit untuk mendukung penerapan pemupukan berimbang. Balai Penelitian Tanah.

Young, H.P. (2009), Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review*, 99 (5), 1899–1924.