Final Evaluation Report: Vietnam Emissions Reduction Challenge Project

Final Report

January 2022

Abt Associates
6130 Executive Blvd
Rockville, MD 20852

In collaboration with:
Denise Mainville Consulting

Suggested citation: Geyer, Judy, Adi Greif, and Denise Mainville. 2022. *Final Evaluation Report: Vietnam Emissions Reduction Challenge Project.* Rockville, Maryland: Abt Associates. Prepared for the Foreign, Commonwealth & Development Office (FCDO) and AgResults Steering Committee.

Submitted to:
Foreign, Commonwealth & Development Office (FCDO)
Abercrombie House, Eaglesham Road
East Kilbride, Glasgow G75 8EA

Submitted by:

Abt Associates
6130 Executive Boulevard
Rockville, MD 20852
Tel: (301) 347-5000
www.abtassociates.com

In partnership with
Denise Mainville Consulting

The AgResults initiative is a partnership between the Australian Government, the Bill & Melinda Gates Foundation, the Government of Canada, the United Kingdom's Foreign, Commonwealth & Development Office, the United States Agency for International Development, and the World Bank.

Table of contents

List of exhibits

Acronyms

AWD	Alternate Wetting and Drying
CO_2e	Carbon dioxide equivalent
DNDC	Denitrification Decomposition
FCDO	Foreign, Commonwealth & Development Office
GHG	Greenhouse gas
GoV	Government of Vietnam
MT	Metric tonne
PfR	Pay-for-results
RCT	Randomised controlled trial
SCP	Structure, Conduct, Performance

Preface

AgResults is a US$152 million multilateral and international learning initiative. It promotes the development and dissemination of high-impact agricultural innovations for food security, health, and nutrition through the design and implementation of prize competitions that are a class of ‘pay-for-results’ (PfR) project. AgResults calls upon the ingenuity and drive of the private sector to both achieve the private sector’s commercial objectives and the donors’ development objectives. AgResults aims to achieve these dual objectives by providing incentives to private sector actors to develop and facilitate the uptake of innovative technologies. The AgResults incentives eventually allow the private sector to overcome market failures impeding the establishment of sustainable commercial markets for these technologies or goods they produce. The AgResults PfR approach aims to achieve substantial and sustained development impacts, including improved food security and food safety, increased farmer incomes, and better health.

AgResults is funded by the governments of Australia, Canada, the United Kingdom, and the United States, and by the Bill & Melinda Gates Foundation. The funds are managed through a Financial Intermediary Fund operated by the World Bank as Trustee. The AgResults team comprises the Steering Committee, Secretariat, Trustee, country-specific Project Managers, and the External Evaluators. The Steering Committee oversees the implementation of AgResults and is composed of the five donors and the Trustee. The Steering Committee is responsible for strategic oversight of the initiative, including endorsement of key management decisions, approval of concepts and business plans for proposed projects, and monitoring of projects and the initiative as a whole. The Trustee provides financial intermediary services. The Secretariat is responsible for implementing the initiative and reports to the Steering Committee. The Secretariat, Monitor Deloitte, contracted SNV Netherlands Development Organisation to manage the project implementation in Vietnam. The Secretariat contracted with Applied GeoSolutions to serve as the third-party Verifier of emissions and yield outcomes.

FCDO, acting on behalf of the Steering Committee, appointed Abt Associates to serve as External Evaluator for six AgResults projects around the world. The evaluator’s role is to use rigorous scientific tools to determine to what extent the prize competitions achieve their objectives to produce private sector behaviours and social outcomes different from, and better than, what would have happened in the absence of the AgResults initiative. The evaluator’s role is vital to the AgResults learning agenda of understanding how donors may leverage the private sector to develop and spread agricultural innovation.

This report presents the evaluator’s assessment of the AgResults Vietnam Emissions Reduction Challenge Project. Judy Geyer, PhD, was Research Director. Denise Mainville, PhD, led the qualitative research, and Adi Greif, PhD, managed data collection for the diaries and income survey and supported quantitative analysis. Phan Thi Ngoc Diep, PhD, served as a local consultant, conducted the baseline commune survey and the farmer diary data collection, and served as a contact point with cooperative leaders. The Mekong Delta Research Institute conducted the farmer survey. Hai Ninh Nguyen Thi, PhD, served as a local consultant, coordinating and conducting qualitative interviews in conjunction with Dr Mainville. Cris Price, ScM, served as quality assurance officer. Diane Ferguson served as editor, and Sally Cameron served as the External Evaluator’s Project Director.

Executive summary

Background

In the Thai Binh province of Vietnam, the AgResults Vietnam Challenge Project (2017-2021) used a pay-for-results (PfR) prize competition to spur development and dissemination of rice production technology packages that both reduced greenhouse gas (GHG) emissions and improved yields. Competitors were private sector companies and non-profit organisations. Each participating competitor developed a unique technology package that combined existing practices associated with GHG reductions, such as alternate wetting and drying and use of short-duration rice varieties. The efficacy of these technology packages was evaluated during Phase 1 of the competition (2017-2019), and competitors with the best-performing technologies were invited to participate in Phase 2, which focused on dissemination. In Phase 2 (2019-2021), competitors pursued seasonal and end-of-project cash prizes weighted according to the number of farmers using the technologies (40%), GHG reductions (20%), yield improvements (20%), and repeat use by farmers (20%).

This report by the External Evaluator evaluates the overall impact of the Vietnam project, with particular emphasis on Phase 2. A separate report (Abt Associates 2019) describes the evaluation of Phase 1. Findings of this evaluation contribute to our learning about the use of PfR competitions to reduce agriculture-induced GHG emissions.

Methods and data

We used a mixed-methods approach to answer a set of evaluation questions common across all AgResults challenge project evaluations. In Vietnam, this approach included:

- Pre-post qualitative comparison to assess the project's impact on private sector involvement in the dissemination of technology packages among farmers
- A randomised controlled trial to estimate the project's impact on farmers' technology package uptake using data from farmer diaries
- A matched comparison design contrasting AgResults farmers to farmers who did not use the technology packages to estimate the impact of technology package adoption on farmer income and yields using data from a large-scale farmer survey
- A descriptive study and cost-effectiveness analysis of program data from the Secretariat and Verifier, including emissions data.

Using the findings from our approach, we also considered the likely sustainability of farmers' use of and competitors' promotion of the technology packages after the project ends. Finally, we considered the project's cost-effectiveness and lessons learnt from implementing a PfR project to develop and disseminate emissions-reducing technology packages for rice production.

Key findings

Private sector involvement: The project succeeded in engaging the private sector in developing and disseminating GHG emissions-reducing technology packages. Of the initial 11 competitors that participated in Phase 1, six companies qualified for Phase 2, of which four chose to participate. The four competitors pursued diverse strategies to promote uptake of their respective technology packages.

Adoption: There is strong evidence that farmers collaborating with AgResults competitors adopted new practices and technology packages. AgResults farmers were more likely than comparison farmers to reduce planting densities and fertiliser use. AgResults farmers were also more likely to use improved crop residue management. Their water management practices were not significantly different from comparison farmers' practices over the whole year, but this result is mostly due to difficulty draining fields during the rainy season. In the dry season, AgResults farmers used less water than comparison farmers.

Yield, income, and emissions: AgResults farmers significantly increased their yields and incomes. Their yields increased by 14% over matched comparison farmers. AgResults farmers had net harvest values (value of production less expenditures) that were, on average, 11% higher than for comparison farmers, on account of higher yields and discounted inputs.

There is not a strong evidence base for assessing GHG emissions. Based on our analysis, both of the data provided by the Verifier to the Secretariat and of data that we commissioned from the Verifier, emissions estimates are highly uncertain. That said, the Verifier's best estimate is a 3–10% reduction in emissions among AgResults farmers, depending on season and competitor. This reduction is smaller than the goal of 30% reduction in the project's business plan.

Sustainability: Two of the four competitors continued to invest in disseminating their technologies in the first season following the competition. Both are very likely to continue to invest in the dissemination of their technology packages given the alignment between the technology packages and their core business models. Farmers were favourable about continuing to use the technology package, although many emphasised that continued use of the packages depended on continuing engagement with the competitors that supported their use of the packages during the project.

Cost-effectiveness: Given the uncertainty of the emissions reduction measurements, we do not report a specific cost per metric tonne of CO_2e reduced. In terms of per-farmer and per-hectare costs to deliver the technology, this PfR project had costs similar to non-PfR projects in Vietnam ($81 per farmer and $747 per hectare).

Lessons about the design and implementation of agricultural prize competitions

Private sector involvement. The AgResults Vietnam Challenge Project demonstrated that PfR approaches can be successful at spurring the private sector to address climate change. In Thai Binh, AgResults competitors increased the supply of emissions-reducing inputs and technology packages, the number of farmers using the technology packages, and the availability of rice produced using these packages. The prize competition did not provide direct incentives to develop markets for the technology packages or their derived goods (carbon credits, rice valued for attributes that are produced using technology packages). Regardless, the two best-performing competitors, for which dissemination of the packages was well-aligned with their core business models, had success in disseminating the technology packages by independently leveraging existing market linkages.

Public-private sector collaboration. This project demonstrated that PfR projects can motivate the private sector to affect public sector action. Like many climate initiatives, GHG emissions reduction in northern Vietnam required action and leadership from the public sector, particularly with respect to water management. A particular synergy may come from PfR public–private sector collaboration compared to direct investment in public sector works and services. To understand this synergy, it could be fruitful to compare the AgResults project to the World Bank-funded One Must Do, Five Reductions program in Vietnam, once it releases GHG emissions information (Jackson et al., 2015).

Verification. Prize sponsors and competitors need to be satisfied that verification procedures are valid and reliable. This is especially important to ensure reliability of estimates of main outcome measures in a development phase before continuing to the dissemination phase. Competitors may also need time during or immediately after the development phase to adjust their innovations to account for emerging results and lessons learnt. The verifier estimating GHG emissions should also apply best practices related to sample size and consistency both in types of measurement over time and in testing under similar conditions multiple times (i.e., in the same season).

Prize structure. Future prize competitions should carefully consider how to best incentivise achievement of the main development outcome. The AgResults Vietnam Challenge Project's prize structure, where winners received prizes proportionate to their results, allowed sponsors to promote multiple outcomes (yield, emissions, farmers reached), and also to strategically put more emphasis, or weight, on specific measures. The grand prize winner ultimately won by disseminating its technology package to large numbers of repeat farmers, but it is unclear whether its technology package substantially reduced emissions. According to the Verifier's emissions estimates, a different competitor had much more substantial emissions reductions and similar yield outcomes but did not win the prize competition.

Theory of change. In the future, emphasising alignment of technology packages with market opportunities could help to increase uptake and sustainability of emissions-reducing technology packages. Future prize competitions should consider awarding prizes for engaging in markets that would incentivise sustained use of the new technology packages. In the Vietnam case these could be markets for emissions-reducing inputs such as slow-release fertiliser, markets for rice produced using emission-reducing technology, and/or carbon offset markets.

Conclusion

The project achieved substantial gains in all areas except GHG emissions reductions, where the results are uncertain. The project demonstrated that PfR can spur substantial private investment in the development and dissemination of high-impact agricultural technologies. This project is one of the first emissions reduction projects conducted with large numbers of smallholder farmers. It is a crucial step in generating greater knowledge of agricultural emissions reduction measurement in the real world; thus, it has lessons for future projects. For future work done on a large scale, improvements in GHG emissions measurement—especially measurements that rely on low-cost indirect observation to facilitate data collection on a large scale—are necessary.

1. Overview of the AgResults Vietnam Emissions Reduction Challenge Project

1.1 Background and challenge project objectives

Rice farming leads to an estimated 7% to 12% of global methane emissions as well as causing even more potent nitrous oxide emissions, making rice a major contributor to climate change (EDF, 2018). In Vietnam, rice is by far the primary contributor to agriculture-based GHG emissions, comprising 48% in 2012 (FAOSTAT, 2017). Reducing GHG emissions is difficult, as GHGs (like other forms of environmental pollution) are a classic example of a 'negative externality', where consumers and producers do not pay the full cost of the pollution, making it difficult to motivate them to incur costs to reduce the pollution. Economists and others have offered multiple solutions for this type of negative externality, ranging from Pigouvian taxes, certification schemes, and emissions regulations to carbon markets. They have also extensively debated the merits of public versus private solutions (foundational thinkers in this field include Arrow, 1970, Beckerman, 1972, and Pigou, 1920).

The Government of Vietnam (GoV) has been concerned about these emissions and has developed goals to reduce them. In its 2012 Action Plan, Vietnam's Ministry of Agriculture and Rural Development set a goal to reduce GHG emissions by 20% along with 20% poverty reduction and 20% economic growth by 2020 (Prime Minister of Vietnam, 2012). One way to reduce GHG emissions and poverty at the same time is to promote adoption of technology packages that simultaneously increase smallholder farmers' yields while reducing GHG emissions. Since farmers are not likely to be motivated to change their practices for the sake of emissions alone, it is crucial to link reduced emissions practices with benefits that will accrue directly to them, such as yield or income increases. Notable programs encouraged by the GoV to reduce GHG emissions include:

- The 'Three Reductions, Three Gains' initiative to reduce quantities of seeds, fertiliser, and pesticide and increase productivity, quality, and efficiency
- The 'One Must Do, Five Reductions' initiative to use certified seeds while reducing use of water, fertiliser, seed, pesticide, and post-harvest waste
- The 'System of Rice Intensification,' a multifaceted approach that includes plant, soil, water, and nutrient management (Demont and Rutsaert, 2017; SRI-Rice, 2015).

These and other projects must overcome two main challenges: (1) understanding what technologies can increase yield and/or income while reducing emissions, and (2) scaling them up through training and aligned incentives for local governments, agribusinesses, and farmers.

The AgResults Vietnam Emissions Reduction Challenge Project tested whether a pay-for-results (PfR) prize competition could help overcome the information and coordination challenges in changing rice-farming practices. The AgResults project offered a PfR prize competition in which private sector firms and non-profit organisations ('competitors', in AgResults parlance) competed to promote farmer adoption of emissions-reducing rice-farming practices. Prizes were awarded to firms according to number of farmers engaged, farmers' repeated use of the technology packages, GHG emissions reduction, and yield increase. The project was designed to incentivise the development of rice-farming technologies that reduce GHG emissions and increase yields, and to promote large-scale adoption of these technologies in Thai Binh province. This province is in the Red River Delta region of northern Vietnam, which has a high density of smallholder farmers—approximately 2.5 million farms. It also has the highest concentration of rice farmer cooperatives in the country, enabling the coordination of farmer groups. It is one of the leading rice production provinces of Vietnam (Guan et al., 2018). Although other institutions are involved in local agricultural programming, there were no other major international donor programs in the province at the time. The Steering Committee, Secretariat, and External Evaluator thus posited that any results of the prize competition would be attributable to the prize competition rather than other programs (Abt Associates, 2017).

1.2 Project design and implementation

Several parties were involved in the design and implementation of the AgResults Vietnam project. With oversight from AgResults' multi-donor Steering Committee, the AgResults Secretariat (Deloitte) designed the project and was responsible for its implementation. The Secretariat contracted the responsibility for in-country project management to the SNV Netherlands Development Organisation. In addition, the Secretariat contracted with Applied GeoSolutions (the 'Verifier') to measure emissions and verify competitor outcomes for the determination of prize awards.

AgResults' Vietnam project consisted of two phases, both implemented by the in-country Project Manager with oversight from the Secretariat and the Steering Committee. In Phase 1, 11 'competitors' tested and tailored technology packages to determine whether a package of practices and technologies would reduce GHG emissions while increasing yield during two rice-growing seasons: Summer 2017 and Spring 2018 (see left side of Exhibit 1-1). Six competitors that met the project's performance standards qualified for Phase 2; of these, four qualifying competitors saw it as within their interests to participate in Phase 2. (See Abt Associates (2019) for an assessment of Phase 1).

Exhibit 1-1. Implementation timeline for the Vietnam Emissions Reduction Challenge Project

Phase 1, Crop 1	Phase 1, Crop 2	Summer 2018 - Gap	Phase 2, Crop 1	Phase 2, Crop 2	Phase 2, Crop 3	Phase 2, Crop 4
Summer 2017	Spring 2018		Spring 2019	Summer 2019	Spring 2020	Summer 2020

Phase 2 promoted scale-up of the qualifying technology packages from Phase 1, with cash prizes totalling $1,850,000 being awarded to Phase 2 competitors for performance relating to yield increases, GHG reduction, and uptake and repeat use of the technology packages by farmers. Phase 2 included four rice-growing seasons (see centre and right side of Exhibit 1-1): Spring 2019, Summer 2019, Spring 2020, and Summer 2020, with awards at the end of each production season and a final award at the end of the four seasons.

Exhibit 1-2 presents details of the incentive structure for the project, including Phase 1 and Phase 2 prizes.

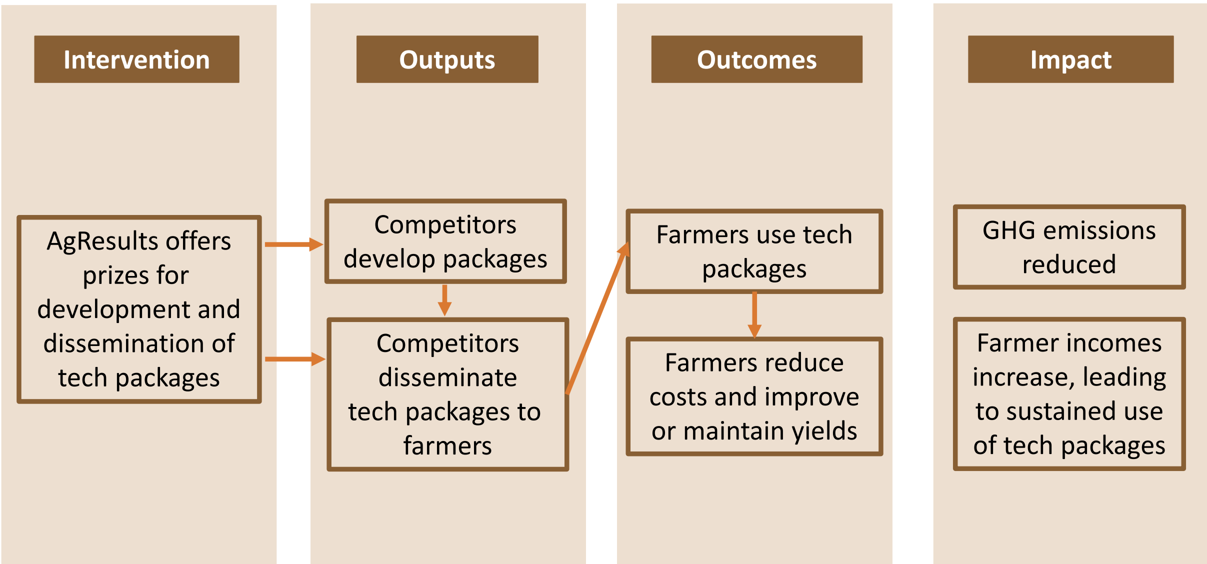
Exhibit 1-2. AgResults Vietnam project incentive structure

TEST				SCALE		
Phase 1 (1.5 years): Competitors test technology packages			→	**Phase 2 (2.5 years): Successful solutions that win Phase 1 are scaled up**		
Interim prize: Competitors who improve upon their baselines for GHG emissions reductions (60% weight) and yield increase (40% weight) share an interim prize of **US$35,000–75,000 proportional to their results**				**Interim prize per crop cycle (3):** Competitors who surpass their competitor-specific baselines for GHG emissions reduction (20% weight), yield increase (20% weight), number of farmers reached (40% weight), and repeat use of tool/product (20% weight) share a prize of **US$500,000 proportional to their results**		
Milestone prize: Competitors with the highest combined GHG emissions reductions (60% weight), and yield increases (40% weight) receive prizes:				**Grand prize**: The three competitors with the highest number of farmers reached (40% weight), repeated use of solutions (20% weight), total GHG emissions reduction (20% weight), and percentage increase in average yield (20% weight) receive added prizes:		
1st Place US$50,000	**2nd Place US$30,000**	**3rd Place US$20,000**		**1st Place US$750,000**	**2nd Place US$400,000**	**3rd Place US$200,000**

Source: Deloitte (2017).

Although the project business plan does not present an explicit theory of change, an implicit theory of change, represented graphically in Exhibit 1-3, is predicated on the assumption that Vietnamese rice farmers will adopt and sustain use of technology packages that improve their well-being through, for example, reduced input costs or increased yields.

Exhibit 1-3. AgResults Vietnam project theory of change



The implicit theory of change linked the AgResults intervention—prizes to competitors for the development and dissemination of yield-increasing and emissions-reducing technology packages—to competitors' efforts to develop the technology packages and demonstrate their effectiveness in Phase 1. The intervention was then expected to lead qualifying competitors to invest in disseminating these packages to farmers in Phase 2. The outcomes of the Phase 2 dissemination included farmer uptake of these technology packages, and the increased yields (and returns) that were expected to result from their application, as well as reduced GHG emissions. These outcomes supported the project's ultimate intended impact: the promotion and sustained use of emissions-reducing technology packages by farmers.

The business plan associated yield increases with sustained uptake by farmers and anticipated that farmers would also benefit from reduced expenditures on inputs. However, yield increases can be associated with either positive or negative impacts on income, depending on the cost incurred in achieving those increases (see Narayan, 2020, for broader discussion). In our evaluation we measure farmer uptake and income effects, accounting for the broader financial aspects of rice production. Suri (2011) notes that benefits and costs of technology uptake explain farmers' technology uptake behaviour. Note that the Vietnam project did not hold the development of markets for the technology or its rice output as either an implicit or explicit goal, or otherwise include it in the project's theory of change.

1.3 Competitors' technology packages

All of the technology packages that the competitors developed included schedules of required inputs and activities to support production of a specific rice variety. Out of ten technology packages, competitors used eight rice varieties. For each rice variety, the competitors developed specific requirements for planting density, flooding and draining schedules, fertiliser use, and rice stubble and straw residue management; together, these practices comprised the technology packages that the competitors were promoting. We discuss each practice in turn.

Planting density. Because methane and nitrous oxide are released through rice stalks, one way to reduce GHG emissions is to have fewer rice stalks, but more tillers of rice from each stalk instead. Doing so can also ensure that yields are maintained or improved while reducing GHG emissions. Planting density can be reduced by transplanting fewer seedlings, or by thinning out stalks by hand after planting (Tivet and Boulakia, 2017).

Water management. Flooding and draining schedules following alternate wetting and drying (AWD) practices reduce the build-up of methane-producing bacteria in flooded fields (Zou et al., 2005; Jain et al., 2013). With training, farmers can use appropriate AWD schedules so that rice plants are flooded at the times the plants require a lot of water for growth and stability as well as to minimise methane and nitrous oxide emissions. AWD can also prevent dilution of pesticides and fertiliser on fields that are not flooded, increasing effectiveness.[1]

Fertiliser use. Reducing the application of nitrogen fertiliser will reduce soil nitrogen and thus nitrous oxide emissions (Liu et al., 2014). Although some amount of nitrogen is necessary to promote rice growth, over-use of nitrogen fertiliser can be prevented by:

- Using 'slow-release' fertilisers. These fertilisers have a coating that wears off in sunlight or water over time, enabling more efficient uptake of nitrogen. These fertilisers are more expensive but require fewer applications.
- Training farmers to understand necessary amounts of nitrogen along with other elements such as phosphorus and potassium that also promote rice growth. In particular, farmers can apply less fertiliser over more applications so it lasts longer (rather than overapplying to account for decreased effectiveness over time).

Crop residue management. Management of rice stubble and straw residue also impacts GHG emissions. Farmers who plant both spring and summer crops need to remove residue prior to the next crop season. The timing between spring and summer crops is too short to allow for natural decomposition over time, leading around 40% of farmers to burn straw and stubble to get rid of it quickly. Burning contributes to carbon dioxide emissions (Truc et al.,

[1] AWD also has unclear consequences for plant protection against disease and pests. Some farmers in qualitative interviews expressed concern that drained fields increased vulnerability to crop damage by rats, for example. However, AWD may be beneficial against other plant diseases.

2012). The application of bioenzymes accelerates decomposition in time to allow for the subsequent season's planting. Decomposition of straw and stubble in the field stores carbon and contributes to yield. Another option is to remove the residue from the field (e.g., by ploughing) to either decompose aerobically or use for other purposes such as mushroom-growing compost or straw brooms (Khosa et al., 2010; Yao et al., 2010).

2. Evaluation overview, analytic methods, and data sources

This section gives a brief overview of the evaluation questions we answer, our mixed-methods evaluation approach, and data sources. Readers who prefer to skip the analysis details may find it sufficient to simply review Exhibit 2-1 and rely on brief methodology summaries in each results chapter. Readers who are interested in more detail will find that Section 2.1 reviews the quantitative methods and data sources used to estimate the impact of the AgResults on technology package uptake and the impact of technology package uptake on farmer income. Section 2.2 describes the qualitative methods and data sources used to analyse private sector engagement in the market, as well as the project's cost-effectiveness and scale.

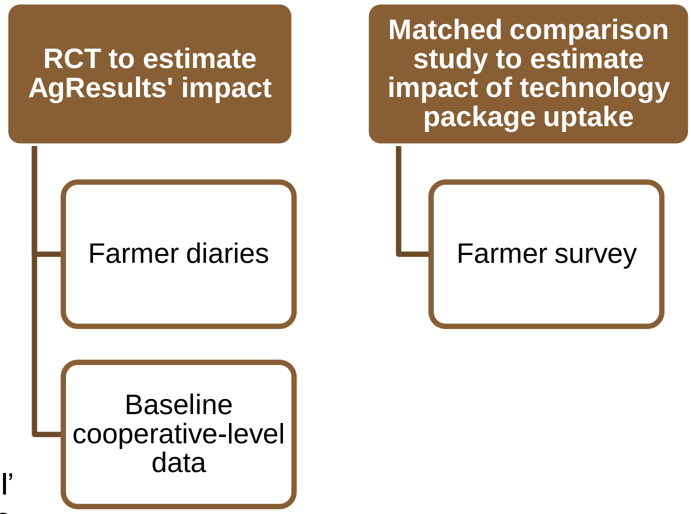
Exhibit 2-1. Evaluation questions and analytic methods

#	Evaluation question	Analytic method
1	**Private sector involvement:** What is the project's impact on private sector involvement in the dissemination of technology packages?	Qualitative approach: pre-post comparison based on the Structure, Conduct, Performance (SCP) conceptual framework guiding key informant interviews and document reviews, integrating findings from Evaluation Questions 2 and 3.
2	**Technology adoption:** What is the project's impact on farmers' uptake of technology packages?	Quantitative approach: (1) randomised controlled trial (RCT) to estimate the AgResults project's impact on average rice cultivation practices in Thai Binh province; (2) quasi-experimental design using matched comparison analysis to estimate impact of being a verified AgResults farmer on farming practices, yield, and GHG emissions.
3	**Farmer impact:** What is the impact of technology package uptake on farmers' incomes, yield, and GHG emissions?	Quantitative approach: Matched comparison analysis to estimate impact of being a verified AgResults farmer on revenue outcomes.
4	**Sustainability:** What evidence exists that the effects of the project will be sustainable in the medium to long term?	Qualitative approach: SCP and qualitative farmer interviews.
5	**Cost-effectiveness and scale:** What is the evidence on the scale of any effect on private sector investment and uptake, and on the cost-effectiveness of the project as an approach?	Qualitative and quantitative approach: SCP and a per-unit cost-effectiveness analysis of farmers and hectares reached.
6	**Lessons:** What lessons can be learnt about best practices in the design and implementation of PfR initiatives?	Synthesis of results from Evaluation Questions 1–5.

2.1 Quantitative methods and data sources

We used different quantitative methods depending on the evaluation question. To answer Evaluation Question 2 on technology adoption, we conducted an RCT within Thai Binh province. To answer Evaluation Question 3 on impacts, we used a matched comparison of AgResults and similar non-AgResults farmers. The selection of comparison farmers is based on characteristics we can observe as similar to AgResults farmers. Without using an RCT, there is always a risk of characteristics that we did not account for (selection bias); we reduced this problem to the extent possible by gathering extensive information from competitors on their selection process and mimicking it.

2.1.1 RCT to estimate the impact of AgResults



The evaluation team with the agreement and help of the AgResults Secretariat and Project Manager embedded an RCT in the prize structure as a randomised incentive design. The RCT allows us to assess the impact of the project on technology uptake by comparing communes in which the AgResults payments were conditionally offered, to communes where no AgResults payments were offered. In July 2018 the evaluation team randomised 50 communes, or 17% of communes in Thai Binh province, to be in the 'control' group. Technology dissemination in the control communes was not eligible for PfR prize payment, as it was in most communes (83%) of Thai Binh province.[2] The 'treatment' communes (N=205), like the control communes, were randomised to their evaluation assignment.[3] Because assignment to the treatment and control group was based on chance alone, we attribute differences between these groups to the impact of the AgResults project. The randomised prize restriction was feasible given the verification protocol, which used geo-referencing to gather yield and cultivation data on individual farmers with whom competitors engaged.

To estimate the impact of AgResults on uptake of rice cultivation technologies, we compared average outcomes of randomly sampled farmers who sell rice, live in the province throughout the rice-growing season, and have high enough literacy and numeracy to fill out the diary detailing their farming practices.[4] Outcome data were collected in calendar year 2020, reflecting farmer outcomes in the second half of the project. We used regression methods to control commune-level baseline covariates, handle commune-level inter-cluster

[2] The analysis strictly uses the random assignment outcomes even though AgResults operated in a limited set of the treatment group communes and there was a small amount of spill-over in the control group. At the onset of Phase 2, competitors appealed the random assignment of three communes to the control group (out of 50). Their appeal was based on their having established a relationship with, or having conducted successful reconnaissance and recruitment of, these three communes for the purpose of AgResults, but before learning about the randomisation. To minimize the evaluation's interference with competitors' plans, these three communes were admitted into the set of communes in which competitors' actions were eligible for prize money. The evaluation team considers these 'spill-over' communes and, in keeping with evaluation protocols, still considers those communes to be in the control group. The evaluation team conducted sensitivity analyses of all contrasts to see how impact estimates change if these three communes are omitted from the analysis: the sensitivity analysis shows that our main findings are robust to this small proportion of 'spill-over' communes.

[3] An additional 31 communes in the province were also eligible for competitor interventions towards prize money, but are excluded from the randomisation study because they either are not rice-intensive (too close to the seashore or too urban) or because they participated in Phase 1 of AgResults.

[4] In the treatment communes where AgResults competitors operated, farmers known to the Verifier as using AgResults technologies were sampled disproportionately (at a higher sampling probability). There were still very few respondents in AgResults – less than 20 per season. The respondents using AgResults technologies were weighted down to make the sample representative of the average within each commune (Annex A).

correlation, and re-weight the data so that the data are representative of all rice farmers in Thai Binh. The covariates and analysis weights are described in Annex A.

Data source: Farmer diaries. In February and March of 2020, we recruited two farmers from each of 262 cooperatives (out of 284 in Thai Binh) to record their rice cultivation practices in a diary.[5] Cooperatives are farmer associations, and there is usually one (but sometimes two) per commune. We randomly selected farmers from lists provided by cooperative leaders of farmers who sell rice, live in the province, and have high enough literacy and numeracy to fill out the diary. We were concerned that the two farmers within each cooperative might have the same water management practices if they use the same drainage pumps and so would not provide different information. To reduce correlation between farmer responses, we coordinated with local cooperative leaders to ensure that the two farmers in each cooperative were served by different drainage pumps.

Local consultants trained farmers on how to fill out the diaries, checked on farmers monthly, and provided ongoing technical support for data collection questions. In addition to their own reviews, consultants also engaged district extension staff to review all diaries and disregarded a small subsample of diaries that farmers filled out incompletely or without adequate detail.

In addition to supporting the RCT, the diary study comprehensively maps current cultivation practices in Thai Binh and how they differ from the AgResults technology packages. It provides a detailed complement on sources of farmer costs to the income survey.

Data source: Baseline cooperative-level data. Both the RCT and the matched comparison study use baseline cooperative-level data to assess the pre-AgResults similarity of communes affected or not affected by AgResults, and to control for any differences in these baseline characteristics while estimating impacts of AgResults. The evaluation team gathered administrative data from nearly all rice farmer cooperatives in Thai Binh province just before Phase 2 began. From October to December 2018, the evaluation team held workshops with 309 cooperative leaders covering 275 communes in Thai Binh to gather key characteristics that the cooperatives keep in their administrative records to monitor their members.[6] The cooperative-level administrative data were from the Spring 2018 rice crop and included details on:

- *Farmer information*: list of all farmers belonging to their cooperative.
- *Rice production*: *cooperative-level* detail on land area under rice, total rice yield and production, the average price per metric tonne (MT) of rice; *farmer-level* information on cultivated rice area, the number of plots, and the elevation of their plots (if known).
- *Irrigation and asset ownership*: whether cooperatives had a completed irrigation system and owned a rice transplanter.

[5] Farmers recorded rice varieties planted; density of rice planted; the dates and water levels at each drain and flood event; the dates of plant maturity (tillering, flowering, harvesting); and details of all inputs with respect to fertiliser, herbicide, and pesticide application. All participants attended a two-hour training on how to record their activity in the diary and received contact information for the Survey Manager, who answered their clarification questions throughout the season. Cooperative leaders also kept a diary of the practices they were recommending to farmers, and the drain and flood actions they took to manage their cooperative's three largest drains (largest built size).

[6] The majority of communes have only one cooperative, but some have more than one.

- *Shocks*: information on disease or weather problems the cooperative experienced.
- *Relationships*: cooperative leaders' knowledge of their farmers' relationships with any of the competitors.

2.1.2 Matched comparison analysis to estimate the impact of technology package uptake

We also conducted a matched comparison study that estimated the rice cultivation practice, yield, and income impact on farmers of working with an AgResults competitor. As AgResults operated in only a very small area and only with a very small proportion of rice farmers, we chose this design to better understand the experience of farmers using the AgResults competitors' technology packages. AgResults competitors worked with 54 cooperative leaders in 51 communes. These were geographically dispersed throughout all eight districts and the province.

To understand farmers' experience, we surveyed over 1000 farmers listed as working with an AgResults competitor. Their responses provide insight on whether the AgResults farmers had a net income loss or gain in each season. To estimate whether the net income loss or gain is attributable to the technology packages, we needed to understand what their net income would have been in the absence of the AgResults project. To achieve this, we surveyed over 1000 matched comparison farmers. The matched comparison study's impact findings are valid for and representative of 85% of the communes in Thai Binh, excluding the 15% of communes that either have little rice production area or an extremely high concentration of rice production area. Annex A provides a detailed overview of the matching technique, but the basic idea is that we assigned larger analysis weights to comparison farmers whose baseline characteristics were most like those of the AgResults farmers. As for any study apart from RCTs, this approach is valid conditional on our assumption that the variables we use to match are the appropriate ones to use. In addition, we also assume that linear regressions are the right way to express the relationship between variables and that the comparison and treatment farmers are mostly similar (have a large 'common support'), so we are not comparing a select subset of treatment farmers to a select subset of similar comparison farmers.

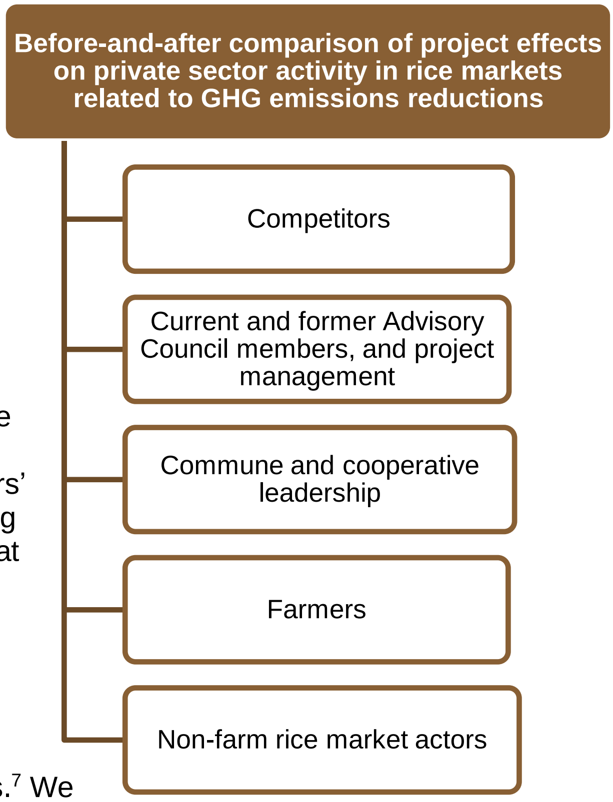
We use a baseline balance test to determine which variables make sense to use. We also do not throw out 'bad matches', but instead use weights on the entire dataset. We do not use any plot-level variables to avoid plot-level selection issues, although we do not believe that farmers selected their best plots to use for AgResults. Farmers had no choice over plots to use, as plot selection was directed by the cooperative leader and had more to do with how that plot related to nearby plots than the individual characteristics (whether they were all high and flat together, rather than one flat plot near a less flat one). We also asked all farmers to list their three best-drained plots. This helped ensure we selected comparison farmer plots most likely to be similar to the AgResults plots.

Despite our best efforts, it is important to recall the limitations of non-experimental designs. A fundamental feature of such designs is that it is impossible to know whether selection bias has been eliminated. Despite this possibility, we find that the results (presented in Sections 4 and 5) are consistent with those of similar projects and are reasonable according to experts the evaluation team consulted.

Data source: Farmer survey. For both the treatment and comparison groups, we recruited smallholder farmers to respond to questions in a detailed agricultural survey. To select respondents, we used two-stage sampling, first by village and then by farmer. Cooperatives are typically comprised of farmers from several villages. AgResults competitors worked through cooperative leaders to recruit farmer groups from specific villages, and we mimicked that selection by fielding our survey at the village level instead of the cooperative level. All farmers, who gave consent to be interviewed and understood that participation was completely voluntary, responded to questions about household demographics, agricultural income, and detailed plot-level information about rice cultivation, input use, harvest, sales, and gendered labour allocation and decision-making. We were fortunate that COVID-19 did not impact survey recruitment due to the immediate and intensive quarantine efforts in Vietnam at the outset of the pandemic. For the summer crop, the survey included 2174 farmers (1090 treatment, 1084 comparison) in 88 cooperatives and more than 180 villages; for the spring it included 2201 farmers (945 treatment, 1223 comparison) in 80 cooperatives and 165 villages.

2.2 Qualitative methods and data sources

The evaluation also relied heavily on qualitative research methods, particularly to frame our investigation of private sector involvement in the dissemination of GHG-reducing technology packages and its sustainability, and to enrich our understanding of the results of Evaluation Questions 2 and 3.



Before-and-after comparison of project effects on private sector activity in rice markets. At baseline and endline, we conducted a qualitative assessment of the rice value chain, focusing on potential markets for emissions-reducing inputs or products, farmers' and firms' perceptions and decisions regarding engaging in those markets, and the effects that AgResults had on those decisions. The assessment drew inspiration from the Structure, Conduct, Performance framework, which is a theory-based, primarily qualitative, approach to market systems analysis that provides a structure to guide the researcher's inquiries into the project's market-level effects.[7] We began by observing the relevant market context in which AgResults competitors sought to scale up use of their respective technology packages, then identifying essential elements of their scaling strategies. We next considered the implications of their scaling strategies within the larger rice market context, focusing on market segments most relevant to reducing emissions. Finally, we assessed the project's performance in motivating private sector involvement in scaling up technology package use among farmers, including sustainability.

The evaluation team conducted semi-structured interviews of project and rice market system stakeholders, seeking insight into respondents' perceptions of the markets' underlying conditions, firms' strategic decisions, and the market's structure and performance. We also drew on project, Verifier, and publicly available secondary data to inform our analysis of the

[7] The SCP framework, detailed in the design report (Abt Associates, 2017), begins by identifying how the underlying characteristics of supply, demand, and the enabling environment of a market lead to the strategic decisions that market actors including firms, farmers, and consumers make about engaging in the market. The strategic decisions of numerous market actors give rise to the market structure, which includes the numbers and characteristics of market participants, the predominant marketing channels, and modes of product transformation and value addition. Together, these factors affect the performance of the market, including whether a sustainable market for the technology develops, how inclusive the market is, and the degree to which it helps meet the objectives that drove the intervention.

markets' structure and performance, and to help triangulate evidence informing our results. The national-level interviews at baseline and endline were led by the evaluation team's qualitative lead, in collaboration with the Vietnam-based agricultural economics consultant. Field interviews were conducted by a team of qualitative interviewers supervised by the Vietnam-based agricultural economics consultant. Exhibit 2-2 displays the number of respondents whom we interviewed.

Data source: Competitors. We collected data from the four Phase 2 competitors about their:

- Phase 2 plans (at baseline), activities, and results (at endline)
- Perceptions of and experience participating in the project
- Perceptions of and experience with segments of the rice market relevant to emissions-reducing technology packages
- Plans to engage with technology packages and rice markets following AgResults' conclusion.

Data source: Current and former Advisory Council members, and project management. We interviewed seven respondents (each at baseline and endline) representing the project's Advisory Council, the Verifier, and the Project Manager to learn about their perceptions of:

- Competitors' Phase 2 activities and results
- Competitors' technology packages and scale-up strategies
- Lessons on the design and implementation of Phase 2 of the project
- The sustainability of the project's impact.

Data source: Commune and cooperative leadership. We interviewed commune and cooperative leaders in both treatment and control communes. At baseline we visited communes where competitors had their treatment plot and a commune where they anticipated scaling their activities in Phase 2. At endline, treatment communes were communes where each competitor worked during Phase 2. Our 'control' communes included communes in Thai Binh where no competitor was active (four at baseline and two at endline), as well as two communes each at baseline and endline in neighbouring Nam Dinh province, where a 'push' project promoted emissions-reducing systems until 2018. The commune and cooperative leader interviews inquired about previous experience, current activities with, and perceptions of technology packages and their components, as well as experience with different rice market segments that could potentially value rice produced with emissions-reducing technology packages. For communes with cooperatives where a competitor was active, we also asked about the process by which the cooperatives were recruited; how they made their decision about whether to participate; and how they identified, recruited, and supported farmers to work with the competitors.

Data source: Farmers. We interviewed four farmers in each of the communes whose leadership we interviewed (we interviewed a total of eight farmers in a commune that hosted two competitors). We attempted to split our farmer sample evenly between male and female farmers, and also between farmers who worked with AgResults competitors (in communes where competitors were active) and those who did not. The interviews explored the farmers' experience with emissions-reducing technology packages and their components and with markets that might value the rice they produced. The interviews also explored intra-household decision-making dynamics around whether to work with AgResults competitors.

Data source: Non-farm rice market actors. We also interviewed a selection of non-farm rice market actors. At baseline, these included traders, processors, and exporters of commodity and specialty rice. At endline, COVID-19 travel restrictions curtailed opportunities for in-person interviews, limiting the overall number of respondents interviewed. (Rice value chain actors that were not affiliated with the project were reluctant to discuss their business activities over the phone with unknown interviewers.) Our endline sample included two rice traders, selected because they bought rice produced by AgResults farmers in communes where competitors who did not purchase rice were active.

Exhibit 2-2. Qualitative data sample

Respondent group							Sample size, all communities	
National-level sample							**Baseline**	**Endline**
Phase 2 competitors							4	3
Advisory Council members, Project Manager, and Verifier							7	7
Other sector experts							5	2
Traders, processors, and exporters							7	2
	Thai Binh treatment		**Thai Binh control**		**Nam Dinh control**		**Subsample size, all communities**	
	Baseline	**Endline**	**Baseline**	**Endline**	**Baseline**	**Endline**	**Baseline**	**Endline**
Commune sample								
Communes	8	7[a]	4	2	2	2	14	11
Within-commune sample								
Commune leadership	1	1	1	1	1	1	14	11
Cooperative leadership	1	1	1	1	1	1	14	11
Farmers	4	4	4	4	4	4	56 (21 female)	48 (22 female)

[a] The endline sample of treatment communes was seven because two competitors were present in one commune. Accordingly, the total farmer sample was 48 because 8 farmers were interviewed (rather than 4) in the commune that had two competitors.

3. Evaluation Question 1: Impact on private sector involvement

The project succeeded in engaging the private sector in developing and disseminating GHG emissions-reducing technology packages. Of the initial 11 competitors that participated in Phase 1, six companies qualified for Phase 2, of which four chose to participate. In Phase 2, all four participating competitors successfully scaled uptake of their respective technology packages to a large number of farmers over the two-year implementation period. Each competitor contracted either cooperatives or individual farmers to produce rice using their respective packages, and supported that production with training, services, and discounted input prices. Two competitors purchased the rice produced by farmers using their technology packages and incorporated it into their core rice business activities. A third competitor used the competition as a means of developing market share and reputation. The fourth competitor promoted its technology package without developing clear linkages to either an input or output market. No competitor attempted to engage the carbon offset market.

In this chapter we first describe the markets that could be affected by demand for, or use of, the technology packages. We then describe the AgResults competitors' strategies within those markets. We analyse several aspects of their strategies, including their choice of market focus, their choice of technology package, the identification and recruitment of farmers, how they disseminated their technology packages to farmers, and how competitors adapted their technology packages to fit farmer or market needs. We paid particular attention to competitors' integration of their dissemination strategies into other rice value chain activities such as the supply of production inputs to rice farmers or post-farmgate sales of rice. The chapter concludes with an overview of how the competitors' market engagement affected the structure of Thai Binh markets.

3.1 Potential markets for emissions-reducing technology packages

Several markets have potential relevance to the emissions-reducing technology packages developed during Phase 1 of the AgResults competition. These include the markets for:

- Emissions-reducing inputs such as slow-release fertiliser.
- The rice produced using emissions-reducing technology packages including mainstream commodity markets and high-value, quality-differentiated specialty markets. (These markets do not value the emissions-reducing traits of the technology packages specifically, but value other aspects of the packages such as consistent quality or limited use of crop chemicals.)
- Carbon offset markets.

Also, the AgResults incentives directly rewarded competitors for the number of farmers using their production systems, essentially creating a "market for farmers" (using emissions-reducing technology packages), in which AgResults served as the buyer and the individual competitors as suppliers.

At baseline, each Phase 2 competitor was aligned with at least one of the aforementioned markets. The four competitors were a specialty rice exporter (i4), two seed companies (i5 and i18), and one fertiliser company (i23). At baseline, competitor i4 was already promoting elements of his technology package to farmers that he contracted to supply rice for his specialty rice export business. Competitors i5 and i18, both seed companies, also described the AgResults competition as consistent with production practices that aligned with the input needs of their seed business—for example, the careful monitoring of production practices and application of production methods that maximised product quality and yield. Finally, i23 described dissemination of their technology package as a way to demonstrate its effectiveness, which would help pave the way to increasing their market share as a supplier of high-quality rice production inputs.

Most competitors also discussed being motivated to help address climate change as an issue—in some cases describing the personal stake they felt in addressing the issue given Vietnam's particular vulnerability to climate change in low-lying areas. Another important motivation that multiple AgResults competitors described was to demonstrate the effectiveness of their technology package to both local government and local farmers, so that they would gain buy-in (particularly with local government) to support the continued dissemination of their systems following AgResults.

3.2 Competitor strategies

We asked each competitor at baseline and endline about the strategies that they planned, and ultimately used, in AgResults' Phase 2. We report on those strategies here, discussing competitors' market focus, how they recruited farmers and disseminated their production systems, and adaptation of the technology packages themselves during Phase 2.

3.2.1 Competitors' choice of market focus

Between our baseline and endline interviews, we found general consistency among most competitors in terms of the market(s) they focused on. At baseline, each of the competitors articulated a clear alignment between their underlying business models and the dissemination of technology packages. Competitor i4 had already been promoting components of the technology package to his contract farmers. He described AgResults as a way to promote a comprehensive system that would reduce costs and improve yields and quality, while enhancing control over production processes that helps maintain access to demanding export markets. Competitor i5 similarly was already promoting the system components, and anticipated yield benefits and cost reductions for farmers, as well as improvements to farmers' production practices that would serve their underlying business interests as a seed company. (I5 also has a processing facility that they use for rice purchases from farmers that do not meet their seed quality requirements or are in excess of their needs.)

Competitor i18 initially took an approach similar to that of i5, with intent to purchase the product from farmers to be processed for either seed or sale on the commodity market. As Phase 2 progressed, however, i18 had difficulty getting farmers to use its varieties given that the rice produced with those varieties does not have a ready market in Thai Binh, and shifted their strategy as described in Section 3.2.3. Meanwhile, i23 took a unique approach relative to the other competitors (but consistent with their own articulated business interests). Competitor i23 focused on scaling up use of their technology package as a way of demonstrating the effectiveness of their fertiliser products. At baseline they reported that they did not intend to buy rice from farmers who produced it using their technology package because there was a ready market available for the rice. They maintained this approach throughout the project.

3.2.2 Competitors' identification and recruitment of farmers

All four competitors described largely consistent approaches to identifying and recruiting farmers to whom they sought to disseminate their technology packages. Competitors began by visiting districts where they already worked or which they thought would be good potential locations for scale-up of their systems. (I5 was particularly successful in leveraging their existing base of farmers, allowing i5 to scale up quickly and also have high rates of repeat farmers from one crop to another.) The competitors collaborated with district extension leadership to hold informational meetings with commune and cooperative leaders in the district. Competitors held follow-up meetings with commune and cooperative leaders who expressed interest in participating, and, if interest was sustained, worked with the leaders to inform the cooperative members of the opportunity. Once a cooperative agreed to work with a competitor, they would identify a potential contiguous area of land,[8] and then recruit individual farmers who had plots within that area. If an individual farmer declined or was unable to participate, the cooperative leader would help to arrange a temporary exchange of land with another farmer who would cultivate the land for the duration of AgResults. Exhibit 3-1 shows competitors' farmer recruitment and retention results by season over the course of the project. Growth in farmer recruitment from Crop 3 to Crop 4 was especially strong owing to competitors' motivation to continue to earn the prize awards. This growth was likely facilitated by concurrent lessening of COVID-19 pandemic threats.

Exhibit 3-1. Farmer recruitment over time

Competitor	Crop 1 Farmers	Crop 2		Crop 3		Crop 4	
		Farmers	Repeat from crop 1	Farmers	Repeat from crop 2	Farmers	Repeat from crop 3
I4	674	1016	20%	1814	29%	5562	47%
I5	791	3910	98%	3895	87%	5052	96%
I18	1345	1179	26%	1334	26%	5878	60%
I23	1567	1094	47%	1202	74%	5052	96%
Total	**4377**	**7199**	**46%**	**8245**	**67%**	**18,878**	**79%**

Source: Compiled from AgResults Secretariat Vietnam presentation at the Spring 2021 Steering Committee Meeting.

Although we hypothesised at baseline that women and/or women-headed households might have lower rates of participation in AgResults due to their lower farm sizes, more limited access to resources, and lesser market integration, a large proportion of AgResults farmers (73.5%) were women. Section 0 explores gender dimensions of technology uptake, participation in AgResults, and yields.

3.2.3 Dissemination of technology packages

AgResults spurred substantial investment by the AgResults competitors to scale up dissemination of their technology packages. In most cases, Phase 2 competitors had never promoted their technology packages as a whole, but had promoted individual components of the packages among farmers in Thai Binh or elsewhere in Vietnam. Nonetheless, the competitors generally described their technology packages as simple and straightforward for farmers to adopt. They also perceived, however, that many farmers would be hesitant to depart from their traditional practices and would need training and reminders to implement the systems consistently. Accordingly, competitors also described local extension services as important partners in the dissemination process and said that it would be important to convince local extension leadership of the merits of their respective technology packages as much as the farmers themselves.

Competitors usually contracted with the cooperatives, but sometimes with individual farmers, and provided key inputs such as seed and fertiliser as in-kind credit, with repayment following harvest. In some cases, the competitors supported efforts to upgrade the irrigation system (necessary for AWD) or provided machinery, such as rice transplanters, that supported implementation of their technology package.

Competitors usually worked through the cooperatives and local extension to train farmers on the technology packages, although in some cases they provided training directly to farmers. An important part of competitor–farmer relationships for two competitors was the use of buyback guarantees, where the competitors committed to purchasing the rice that farmers

---

[8] The parcel had to be a minimum of 1500 square meters to allow for verification.

produced using their technology package. For i4 and i5, there was a direct incentive to buy the output because they had use for it in their core business activities. Similarly, i18 initially offered to buy back the rice, but later stopped doing this and had farmers dispose of their rice however they desired, often on the commodity market. I18 subsequently incorporated the rice variety BC15 (produced by Thai Binh Seed) into its technology package in order to make the package more acceptable to farmers by making it easier for them to sell the rice they produced. BC15 is commonly grown in Thai Binh and so is easily sold by farmers; whereas the competitor had previously found that its own varieties, which are not commonly grown in Thai Binh, were difficult for farmers to sell. This resulted in farmers being reluctant to adopt i18's technology package because it was difficult for them to find a market for the rice they produced. Only i23, the fertiliser company, did not make any sort of buyback offer to farmers producing rice with its technology package, instead emphasising that the rice could be easily sold on the commodity market.

3.2.4 Adaptation of technology packages

After Phase 1 ended, and throughout Phase 2, the competitors adapted their technology packages to enable their dissemination. As competitors needed to stay within prize competition rules, they had a restricted range of options for adaptation. Two types of adaptation were most common: simplification, and expansion and adaptation to incorporate new rice varieties. Simplification was a strategic activity described by several competitors to make it easier for farmers to implement the packages; for example, by reducing the number of fertiliser applications. Competitors chose to simplify in order to facilitate farmer uptake, even though they anticipated trade-offs with efficiency, yields, or GHG reductions to result. Similarly, some competitors shifted from requiring transplanting to allowing seeding as well.

Adapting packages to allow for new varieties (and in some cases to drop varieties) was another approach that some of the competitors used to increase their scale. Competitors were allowed to add new rice varieties as long as the variety was used in Phase 1 (even if by another competitor). (DS3 was an exception that was allowed in after a field demonstration.)

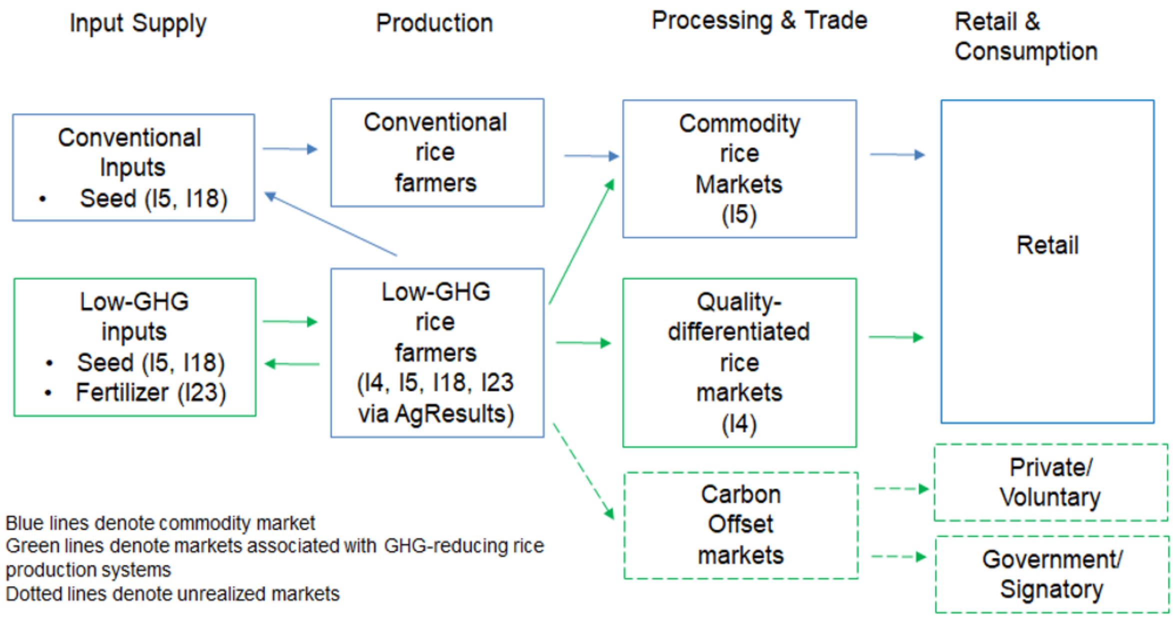
3.3 Rice market structure

AgResults increased the supply of emissions-reducing inputs and technology packages, the number of farmers using emissions-reducing inputs and technology packages, and the availability of rice produced using them. We did not find evidence that AgResults changed the structure of Thai Binh's rice market, nor, by design, was it intended to. Nonetheless, AgResults fostered the creation of competitor-specific market channels. Competitors used these channels to supply farmers with emissions-reducing inputs and technology packages, to take possession of the rice produced using those systems, and then to feed it into their existing market channels.

Exhibit 3-2 is a simplified graphic of Thai Binh's rice market system. In the top tier, it shows the conventional, undifferentiated rice market that dominates in Thai Binh. The bottom of the graphic depicts the rice market system that reflects activity induced by AgResults. Here, inputs for rice produced using emissions-reducing technologies are supplied by the competitors, while the competitors also 'supply' farmers to AgResults. Following production, farmers supply rice output either back to the seed companies (such as i5 and i18) for processing as seed or the commodity market[9] or to high-value rice buyers (such as i4). The dashed lines show that, despite the hypothetical potential for them to do so, competitors did not sell into any carbon offset market during AgResults.

[9] Excess seed or rice not meeting seed requirements can be sold for consumption to commodity (rice) markets.

Exhibit 3-2. Structure of Thai Binh's rice market system



4. Evaluation Question 2: Impact on technology uptake

There is strong evidence that farmers collaborating with AgResults competitors adopted new practices and technology packages. AgResults farmers were more likely than comparison farmers to reduce planting densities and fertiliser use. AgResults farmers were also more likely to use improved crop residue management. Their water management practices were not significantly different from comparison farmers' practices over the whole year, but this result is mostly due to difficulty draining fields during the rainy season. In the dry season, AgResults farmers used less water than comparison farmers. The project verified 28,031 unique farmers using the technology package at least once over four crop seasons. Competitors engaged farmers throughout the province, reaching 6% of the farmer population. Thus, the AgResults competition did not have significant impacts on the average Thai Binh rice farmer's uptake of any technology package developed.

In this section we first use the RCT design and data from the farmer diaries to address the key question "What is the impact of AgResults on the uptake of technology packages?" We also use the RCT design to show current technology practices, which makes it possible to systematically assess the room for projects such as AgResults to make changes. The second half of this section uses the quasi-experimental design and the farmer survey of AgResults farmers and comparison farmers to describe the impact of technology adoption on farmer practices, perceptions, and differences in technology adoption or other equity challenges by gender.

4.1 Impact of AgResults on technology package uptake across Thai Binh

The project reached 28,031 unique farmers across the four seasons of the competition. This scale of adoption is similar to or larger than the scale achieved by similar projects in Vietnam such as the World Bank-funded One Must Do, Five Reductions (Jackson et al., 2015), Lộc Trời/IFC's Pilot & Promote SRP (2015-2018) (World Bank Group, 2016), and GIZ/Olam's 's Better Rice for Asia BRIA II (2018-2022) (BRIA II Factsheet, 2021). The scale is far short of the goal of 75,000 farmers, mostly because there were fewer competitors than expected (the number of farmers per competitor was roughly aligned with expectations). The project reached approximately 3% of Thai Binh's rice production area, or 6% of the rice farmers. It also reached approximately 22% of communes and 20% of cooperative leaders assigned to treatment areas. The treatment farmers were distributed throughout the province without any apparent patterns by location.

Our RCT of the province (across locations assigned to treatment and control), based on diary studies of two randomly selected farmers per cooperative, found that very few farmers in Thai Binh follow any technology package comprehensively, although many applied specific components of the packages. In control communes, this shows the potential for scaling up technology packages. In treatment communes, this means that there was little spill-over of farmers using AgResults to nearby farmers. This small overall scale means we do not expect to see province-level differences in outcomes.

Indeed, our RCT findings confirm the lack of changes in technology component uptake at the province level; see Annex D, Exhibit D-1, for details.

4.2 Differences between the technology packages and standard practice

Data from the RCT illustrate current practices in the province and show that almost no control farmers use all of the technologies together in ways similar to the AgResults technology packages. The AgResults project emphasises that the technology package overall provides synergistic benefits, so the whole is greater than the sum of its parts. That said, it is also useful to look at individual components of recommended packages to understand where farmers are already implementing best practices and where there is room for improvement.

The components of the technology practices are all already practiced, but to varying degrees. Exhibit 4-1 shows the estimated proportion of rice farmers in the control communes who followed the components of the technology packages of each competitor, using data from the farmer diaries. We limit this table to the control communes to avoid any confusion about whether AgResults influenced these observations. Since there is high overlap between some of the competitor technology packages, some comparison farmers are counted as following the recommendations of more than one technology package.

The majority of farmers already follow most recommended best practices, apart from planting density, use of bioenzymes, and days drained. Many farmers already use AgResults-approved rice varieties (40-60%, sometimes higher in the summer), the recommended number of fertiliser applications (approximately 70%) and recommended amounts of nitrogen (approximately 70%). Only 25-36% are within the recommended range for planting density, use of bioenzymes or lime (almost always lime), and days the field is drained completely dry. Although the technology packages did not impose requirements or limitations on burning post-harvest straw and stubble, we note that almost 40% of farmers burn crop residue, which releases GHG emissions (figure not shown in exhibit).

Exhibit 4-1. Use of competitors' component technologies in the control group

4.3 Changes in farmer practices

We also compared the practices of AgResults farmers to those of matched comparison farmers. According to the matched comparison analysis, AgResults farmers had 7% lower planting density, 12% more frequent fertiliser applications, and 13% lower added nitrogen. Differences in water management depended on the season, with AgResults farmers reporting around five more dry days in the spring and two fewer dry days in the summer crop, compared to the matched comparison group. AgResults farmers were also much more likely to use a competitor's recommended rice variety, fertiliser brand, and bioenzymes. Exhibit 4-2 displays the impacts of technology package uptake on these farmer practices (Exhibit D-2 in Annex D shows the results separately by crop season).

To understand the results of AgResults farmers compared to the package requirements, see the first column, which shows the most common technology package requirement across the packages for each outcome/component. The package components have synergistic effects and so are hard to evaluate separately. Also, some competitors do not comply with a practice considered emissions-reducing for one component because they use emissions-reducing practices on another component instead. For example, i5 only requires two fertiliser applications, which is a current standard practice rather than an emissions-reducing one, but asks for reduced planting density and other practices instead to reduce emissions overall.

That said, even though the components interact in complex ways, the more of these components that are used, likely the better the result for GHG emissions reductions. AgResults farmers on average complied with average expectations for nitrogen fertiliser applied (2.6 kgs/sao compared to a recommendation of 3.64 kgs/sao) and number of times applied (three applications). AgResults farmers did not quite reach the average recommendation for planting density (at 1.4 kgs/sao compared to a recommendation of 1.33 kgs/sao). Almost all (90%) of farmers reported using a recommended rice variety, and this may be an underestimate due to farmers calling some rice varieties by multiple names. The average number of dry days in the field (11) did not meet the average package recommendation of roughly 17 days (on average a 14% reduction).[10] Bioenzymes were originally required by almost all of the competitors, and none of them recommended burning, yet we see that not all farmers use bioenzymes, and a substantial minority still burn their fields. However, the Phase 1 testing showed bioenzymes actually did not reduce GHG emissions as expected, and competitors did not always require it. Burning was not banned. Therefore, lack of bioenzyme use and field burning are not against the prize rules.

Exhibit 4-2. Impact of technology package uptake among AgResults farmers

Outcome	Most common AgResults technology package recom-mendation	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	% Change (A-B)/B	Standard error	Signifi-cance (p-value)
Used AgResults-equivalent rice variety (%)	N/A	90.0	52.8	37.2***	70.5	3.3	0.000
Planting density (kgs/sao)	Less than 1.33 kgs/sao	1.4	1.5	-0.1**	-6.7	0.0	0.011
Used fertiliser variety	Varies by competitor	72.5	27.7	44.8***	161.7	2.8	0.000

[10] The recommended number of dry days depends on the rice variety, the planting method (transplanting vs. sowing) and the season. Across competitors' technology packages, there was variation in recommended number of dry days even for the same rice variety, planting method, and season.

Outcome	Most common AgResults technology package recommendation	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	% Change (A-B)/B	Standard error	Significance (p-value)
recommended by a competitor (%)							
Number of times apply fertiliser	3 or more	3.0	2.7	0.3***	11.1	0.1	0.000
Nitrogen applied (kgs/sao)	Less than 3.64 kg/sao	2.6	3.0	-0.4**	-13.3	0.2	0.025
Number of days the plot was completely dry	14% or more, roughly 17 days	11.1	11.9	-0.8	-6.7	0.9	0.410
Used bioenzymes on straw (%)	Varies by competitor	39.0	4.7	34.3***	729.8	3.5	0.000
Burned straw (%)	0	35.20	41.10	-5.9*	-14.4	3.3	0.071

Source: AgResults Independent Evaluator's Farmer Survey.
Note: One 'sao' is equal to 360 square meters; we use 'sao' because the technology packages were communicated to farmers in terms of sao. Means are regression-adjusted. The number of treatment and comparison farmer plots in the spring is 2835 and 3669; in summer, 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

4.4 Farmer perceptions

We complement our evidence on AgResults' impact on uptake with insights into farmer perceptions of the different components of the competitors' technology packages for rice production. As described in our baseline assessment, and supported by our endline results, farmers generally described an array of benefits from applying different components of the technology packages, helping us to understand why they would be motivated to use technology package components without necessarily having any direct incentive to prioritise reducing emissions. Across the array of technology components, these included, variously, improved plant health, rice quality, and yields, and reduced production expenditures and labour requirements. Many farmers interviewed had been using specific practices for a long time—sometimes for 10 years or more—and described having learned of them from training programs offered by the government extension service or agricultural university, sometimes in conjunction with an input supply company. Generally, farmers had limited awareness of climate change issues, the role that rice production played in climate change, and the potential for the specific technology components to reduce emissions.

In many of our qualitative interviews at endline, farmers reported using different technology package components because of the guidance of their cooperative leader, and not forming personal opinions on the merits and drawbacks of the technology components until after using them and seeing their results. While being aware of technology packages is arguably a precondition to using them, our qualitative results suggest that farmers' attitudes about the specific technology packages only evolved or emerged once farmers obtained first-hand experience of them. Positive attitudes towards the individual technology packages are critical to sustaining their use, however.

In addition to qualitative data, we used the income survey to assess farmers' awareness of and attitudes towards the planting density, fertiliser requirements, water management requirements, and crop residue management. Exhibit 4-3 shows that almost all AgResults farmers had heard of the technologies, and few reported any challenges. AgResults farmers do not primarily associate the technology components with GHG emissions reductions, with few reporting lower GHG emissions as a benefit of planting density, fertiliser requirements, or water management. Some (24%) associate crop residue management with lower GHG emissions. The majority (60%) report that the planting density requirements improve yield, whereas nearly a quarter report that fertiliser requirements and drain requirements improve yield.

Exhibit 4-3. AgResults farmers' assessment of benefits of AgResults technology components

Benefit or challenge (top 3 for each component) [a]	Draining plots	Reduced plant density	Crop residue management	Reduced use of fertiliser
Common benefits				
Has heard of this technology	91.8%	98.4%	94.0%	91.8%
Higher yield	23.1%	59.7%	6.9%	21.0%
Lowers GHG emissions	0.8%	0.1%	24.1%	0.8%
Component-specific benefits				
Lodging tolerance [b]	33.5%			
Better quality rice	24.7%		8.5%	26.8%
Less crop loss	19.0%			
Reduced pests		63.8%		
Less labour		41.7%	10.7%	26.0%
Uses fewer inputs		30.0%		31.8%
Better soil quality			76.7%	
Percentage reporting no challenges or did not use this technology				
No challenges	33.5%	92.6%	59.2%	84.2%
Did not use this technology		1.2%		
Component-specific challenges				
Too difficult to drain land	24.7%			
Requires too much coordination with other people	19.0 %			
Reduces yield		1.3%		5.5%
More difficulty in land preparation			15.8%	
Stubble not completely decomposed			11.2%	
Requires too much time/labour				5.1%

Key: Lighter shading corresponds to lower percentages; darker shading corresponds to higher percentages.
Source: AgResults Independent Evaluator's Farmer Income Survey.
[a] Respondents could select up to three responses from a list read to them for each component.
[b] Lodging tolerance = "less likely to bend towards the ground."

4.5 Technology uptake by gender

Our baseline assessment highlighted that women play a prominent role in every aspect of rice production in Thai Binh, yet female-headed households tend to have less land to cultivate and less of a commercial orientation to their production. Consequently, we hypothesised that female-headed households could be disadvantaged in AgResults if competitors focused on engaging farmers with more land or used linkages to high-value markets to increase uptake of technology packages. At the intra-household level, given the *de facto* prevalence of women in rice production, we hypothesised that women could be disadvantaged if the technology packages promoted by AgResults competitors increased labour requirements. Conversely, the promoted packages could advantage women if they decreased labour requirements.

To investigate these hypotheses, we drew on results of our semi-structured farmer interviews and our farmer income survey study to investigate gender patterns underlying participation in AgResults and different rice-related production activities. We also investigated whether there were gender patterns behind important farmer perspectives about the benefits and drawbacks of different technology package components.

Overall, despite finding some gender-differentiated patterns among those results, no clear patterns of advantages or disadvantages in any of these areas of inquiry emerged. (Exhibit D-3 in Annex D displays the differential impacts of AgResults for female- and male-headed households.)

Participation: The majority of AgResults survey respondents from the matched comparison study were women, indicating that female gender was not a hindrance to inclusion. Within the AgResults sample, female-headed households had 23% less area for rice than male-headed households, on average. In our qualitative interviews with farmers, we asked about the process by which the household decided whether to participate in AgResults. Gendered patterns of inclusion or exclusion did not appear to influence that decision.

Production: Our baseline results showed that there were few culturally determined roles for women versus men in rice production, although women, particularly older women, were heavily involved in rice cultivation. Our endline survey reinforced this perception. In particular, women were characterised as being responsible for the bulk of all rice production activities except the application of pesticides and herbicides to the rice plot. They were most likely to be responsible for thinning rice seedlings and for weeding. Among female-headed households, participating in AgResults decreased the likelihood that women would be solely responsible for field preparation and transplanting. A possible explanation for this is that female-headed households working with AgResults competitors may have been more likely to have outside help from neighbours or other farm workers, as part of the effort to collaborate within a composite AgResults site. Participating in AgResults increased the likelihood that women would be solely responsible for drainage management. A possible reason for this role reversal is that the practice of AWD by AgResults-participating farmers requires more frequent management of water levels, leading to a pragmatic reliance on women, who might have greater availability for the task.

Perceived benefits and drawbacks: Our quantitative analysis of AgResults farmers revealed some significant and notable differences in the perceptions of male versus female farmer respondents. Female respondents were more likely than male respondents to report that the low-density requirements reduced labour. They were also less likely to report that the fertiliser requirements reduced inputs, or that draining requirements reduced lodging (the collapse of stems at ground level), or that ploughing stubble reduced soil quality. Taken together, there is no clear pattern suggesting that gender-based differences in perceived benefits and drawbacks would result in gender-based differences in future use of the technology packages.

5. Evaluation Question 3: Impact on income, yield, and GHG emissions

Yield, income, and emissions: *AgResults farmers significantly increased their yields and incomes. Their yields increased by 14% over matched comparison farmers. AgResults farmers had 2% lower expenditures (with discounts) but also sold rice at prices that were on average 6% lower. Taken together, the AgResults farmers had net harvest values (value of production less expenditures) that were, on average, 11% higher than for comparison farmers, on account of higher yields and discounted inputs. Our analysis finds that despite higher yields, AgResults farmers' higher net harvest values would not be sustained if input price discounts were removed, even if they sold rice at the same average prices as comparison farmers.*

There is not a strong evidence base for assessing GHG emissions. Based on our analysis, both of the data provided by the Verifier to the Secretariat and of data that we commissioned from the Verifier, emissions estimates are highly uncertain. That said, the Verifier's best estimate is a 3–10% reduction in emissions among AgResults farmers, depending on season and competitor, for an average reduction of 0.69 MT/ha. This reduction is smaller than the goal of 30% reduction in the project's business plan.

This chapter compares AgResults farmers to the matched comparison group using our 2020 survey of over 2100 farmers. First, we describe the impact of technology uptake on net rice income and revenue. Second, we analyse net rice income in the absence of input discounts and changes to sales prices. Third, we consider yield as an outcome important to understanding the net value results. Fourth, we consider the changes in farming practices that resulted in yield improvements. Fifth, we investigate potential gender differentials in our outcomes of interest. Finally, we use Verifier data to discuss the project's impact on GHG emissions.

5.1 Impact of technology uptake on net rice income

Using three different definitions of rice income, we found that AgResults farmers fared better than farmers who were not involved with AgResults. Exhibit 5-1 displays the net value, gross sales revenue, and net sales revenue for AgResults farmers (considering only the plots on which they used technology packages). It shows the difference compared to matched comparison farmers and whether that difference is statistically significant. Prices are adjusted for comparability based on whether rice was sold dry or fresh and by rice variety. Matching variables are at the household and plot level. Given that participating fields were selected at the cooperative level on the basis of factors such as access to adequate drainage and the availability of adequate adjacent plots to meet minimum field size requirements, we do not anticipate any plot-level selection bias. (See Annex A for more discussion.) The means are regression-adjusted and weighted. Exhibit E-1 in Annex E displays the results by separately by season.

Exhibit 5-1. Impact of technology package uptake on revenue

Outcome	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	% Change (A-B)/B	Standard error	Significance (p-value)
Net value – value if sold all harvested rice at sales price minus expenditures ('000 VND/sao)	887.3	802.3	85.0 **	10.6	35.2	0.017
Gross sales revenue ('000 VND/sao)	898.8	550.5	348.3 ***	63.3	58.2	0.000
Net sales revenue (income minus expenditures) ('000 VND/sao)	270.0	-75.9	345.9 ***	455.7	57.7	0.000
Total expenditures ('000 VND/sao)	568.0	580.3	-12.3	-2.1	12.5	0.326
Rice sale price ('000 VND)	7.5	8.0	-0.5 ***	-6.3	0.1	0.000
Proportion of farmers that sell rice (yes/no)	50.0	38.0	12.0 ***	31.6	2.8	0.000
Amount of rice sold (kgs)	388.1	281.9	106.2 ***	37.7	32.3	0.001

Source: AgResults Independent Evaluator's Farmer Income Survey.
Note: Sales prices are lower on average for AgResults farmers because they are only one element among a bundle of terms that comprised competitors' contractual arrangements with cooperatives or farmers. Contract arrangements often included the provision of discounted inputs, in-kind credit, and services (such as collection of fresh rice from the fields). All of these imply costs borne by the competitor that are not explicitly accounted for in the sale price reported by farmers. Although one competitor provided a large markup in price, they provided an even larger markup in price to the comparison farmers with whom they worked.
Note: Means are regression-adjusted. The number of treatment and comparison plots in the spring is 2835 and 3669; summer 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

Net value. For the study's main income variable of interest, we examined the net value of all the rice on a plot, as if it were all sold, so that we could compare all farmers regardless of whether they sold rice. AgResults farmers' average net value was 11% higher than comparison farmers' (85K VND per sao higher than the comparison group's average of 802K VND per sao). This difference is statistically significant. To find this result, we defined the net value of the rice harvest as the total amount harvested times the sales price of the rice, minus production expenditures. For the sales price, we used the price the farmer got for any sale of the harvest. If the farmer did not sell the rice, we imputed the sales price using information from farmers in the same commune who sold the same rice variety.[11] Net value for AgResults farmers averaged VND 887K per sao per season in 2020, higher than the Project Manager's estimates of VND 500K–630K.[12]

Gross sales revenue. AgResults farmers' gross sales revenue improved significantly. Gross sales revenue is the amount farmers received for what they actually sold. Average

---

[11] We also standardised all prices to the same moisture level. Note that if all farmers switched to selling their entire harvest, the price of rice could drop, but this provides a reasonable estimate of total value.

[12] To check whether our estimates were within a reasonable range, we compared our findings to a 2019 analysis of competitors' expected revenue and expenditures written by the Project Manager using market information and on-site farmer interviews. The Project Manager found that income per sao, net of expenditures, was always expected to be positive, and was on average VND 501K per sao in spring, and VND 636 per sao in summer, not including family labour cost (equivalent to VND 13,918K per hectare in spring and VND 17.8K per hectare in summer). Our findings are similarly large and positive, although with an even higher average income of 881K per sao per season in 2020.

gross sales revenue is much lower for comparison farmers (VND 550K per sao) than for AgResults farmers (VND 898K per sao), consistent with the fact that a smaller proportion of matched comparison farmers sold any rice. On average, 48% of farmers who worked with AgResults competitors sold any rice, compared to 36% of the matched comparison farmers. The higher sales rates are driven by the competitors that bought back rice; farmers working with competitors that did not buy the rice back sold at the same rate as farmers in the comparison group.

Net sales revenue. Net sales revenue, gross sales revenue minus expenditures, was also much higher for AgResults farmers than it was for comparison farmers. Higher net sales revenue is explained by the statistically significant increase in sales. Farmers working with AgResults competitors were, on average, 12 percentage points more likely to sell rice than the matched comparison farmers; they also had more rice to sell, due to their higher yields (see Section 4.4).

Expenditures. In the net value and net revenue calculations, expenditures included major expenses such as land preparation, machinery services, inputs, and labour expenditures.[13] We expected lower expenditures for AgResults farmers due to discounted prices for inputs. We did not find overall differences in average expenditures, likely because many of the high-priced items were offered with price discounts (such as bioenzymes, slow-release fertilisers, or unusual seeds). AgResults participation influenced a relatively small proportion of total expenditures; it does not affect many major expenditures such as land preparation and harvesting fees.

Sales prices are lower on average for AgResults farmers because they are only one element among a bundle of terms that comprised competitors' contractual arrangements with cooperatives or farmers. Contract arrangements often included the provision of discounted inputs, in-kind credit, and services (such as collection of fresh rice from the fields). For the competitors that bought back the rice, they may have reduced the rice sale price to recoup some of the cost of the discounts they had provided for inputs. One competitor provided a 30% price markup to help create a stable, quality-controlled group of farmers that will sell to them. However, the price markup this competitor provides to other similar farmers in the comparison group is even higher.[14]

5.2 Impact of technology uptake on net rice income without discounted input prices

To understand whether the AgResults project would lead to increased revenue even without free or discounted inputs (or rice price markups/markdowns), we analysed expected income without these cost or sales price changes associated with participation in AgResults.[15] We

[13] We also considered but ultimately excluded other expenditures such as cooperative fees and expenditures for rat killing, root stimulants, leaf blight protection, anti-streaking and stem borers, and land rentals. These were too small overall to influence the outcomes and were too variable/defined differently by each cooperative to be comparable. Some of the larger expenses such as land and labour expenditures were not asked about in the AgResults farmers survey but instead are imputed from the diary study, accounting for the different characteristics of diary and AgResults survey farmers. These expenditures did not vary enough to affect outcomes. We also excluded reported prize awards; they were not distributed by the time of the survey.

[14] This conclusion is based on our analysis from the farmer income survey.

[15] In actuality, without AgResults, farmers would not have grown rice in the same way, or all of the same rice varieties, but this is an exercise to understand whether the technologies are sustainable even without discounted input prices for the future. We gathered price discount information from the farmer survey responses, diary responses, semi-structured farmer interviews, selected cooperative leader interviews, and competitors. Benefits received varied by cooperative, even those working for

used input cost information from the comparison farmers to estimate what AgResults-verified farmers would have paid to use the technology packages if they had faced the full (not discounted) input costs and sales prices (without markups or markdowns) as the comparison farmers.[16] Columns A and B of Exhibit 5-2 show the actual revenue outcomes for AgResults farmers and the hypothetical revenue they would have had in the absence of discounted input costs. Data in column B assume AgResults farmers receive the same prices for their rice sales as the comparison farmers (by variety and moisture level). To estimate revenue impacts in the absence of discounted input prices, we compared the expected revenue (column B) with the actual revenue of comparison farmers (column C).

Net value and net sales revenue decrease without discounted inputs but remain positive. In the absence of discounted input prices, average net value per season is not higher for farmers using AgResults competitors' technology packages. Net sales revenue remains significantly higher for AgResults farmers due to higher yields and a higher percentage of farmers who sell. This result is driven by the expected expenditures in the absence of AgResults discounts becoming higher than what is typical. For example, we observe that AgResults farmers paid reduced prices ranging from 30K to 37K VND/sao in spring for BC15, compared to 42K VND/sao in the comparison group (Annex E, Exhibit E-2). The reported fertiliser expenditures for the competitors' technology packages are highest for i4's technology package for DS1 (Japonica rice) but otherwise comparable and in the range of the fertiliser expenditures of the comparison group. For results by technology package see Annex E, Exhibit E-3 and Exhibit E-4.

Exhibit 5-2. Imputed impact of technology package uptake on revenue in the absence of AgResults-associated competitor incentives

Outcome	Mean of farmers' AgResults plots (A)	Imputed mean in absence of competitor incentives (B)	Mean of all comparison farmer plots (C)	Difference (B-C)	% Change (B-C)/C	P Value
Net value ('000 VND/sao)	887.3	828.6	802.3	26.3	3.3	0.329
Net sales revenue (income minus expenditures) ('000 VND/sao)	270.0	198.6	-75.9	274.5 ***	361.7	0.000
Total expenditures ('000 VND/sao)	568.0	608.1	580.3	27.8 **	4.8	0.024

Source: AgResults Independent Evaluator's Farmer Income Survey.
Note: Means are regression-adjusted. The number of treatment and comparison plots in the spring is 2835 and 3669; summer 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

5.3 Impact of technology uptake on yield

Yield is a key outcome in the theory of change and is integral to net value calculations. We found that average yields for AgResults farmers were 5.6 MT/ha and that the plots where competitors' technology packages were applied had 14% higher yields compared to the matched comparison group. The difference in yield is statistically significant. Exhibit 5-3

---

the same competitor, but we did see some general patterns. One competitor discounted seed prices, another allowed farmers to buy seeds on credit to be paid at time of harvest. All competitors provided in-kind discounts on fertiliser. Some competitors provided in-kind subsidies of other inputs, such as bioenzymes.

[16] By comparing information from AgResults and comparison farmers, we account for reduced input prices that farmers might receive even without AgResults (cooperative-provided fertiliser was the most common).

displays the mean yield on plots where AgResults competitors' technology packages were applied, the mean yield of comparison farmers' plots, the difference, and the statistical significance. The mean yield increase is 0.7 MT/ha per season, which is modest compared to expectations based on Phase 1 estimates of the technology packages' performance. These expectations were increases of 1.8, 3.0, 5.4, and 1.0 MT/ha per season for i4, i5, i18, and i23, respectively, which are ambitious given that an increase of 1.5 MT/ha corresponds to a 30% increase.

Exhibit 5-3. Impact of technology package uptake on yield

Outcome	AgResults plots of AgResults farmers (A)	Farm plots of comparison farmers (B)	Difference (A-B)	% Change (A-B)/B	Standard error	Significance (p-value)
Yield (metric tons/hectare)	5.6	4.9	0.7 ***	14.3	0.1	0.000
Amount of rice harvested (kgs/plot)	764.6	681.2	83.4 **	12.2	37.8	0.029
Area of plot (sao)	3.4	3.2	0.2	6.3	0.2	0.387

Source: AgResults Independent Evaluator's Farmer Income Survey.
Note: Means are regression-adjusted. The number of treatment and comparison plots for which we have information on yield in the spring is 2082 and 3669; summer 2319 and 3270. These numbers vary slightly by outcome due to sporadic missing data (for example, farmer does not respond to that question in the survey).
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

The increase in yield (14%) is roughly comparable to the increase in net value (11%). The fact that they are not equal is likely due to a combination of the lower average sales price in the treatment group, and random measurement error. Because the statistical precision is limited, the confidence intervals of the impacts overlap: the data do not reject the hypothesis that yield and net value increased at the same levels.

Our data on yields suggest higher yield impacts than those from the Verifier. For the same crop seasons that we studied, the Verifier reported an average yield increase of 1% in spring and a decrease of 1% in the summer. The Verifier's estimate may be lower partly because the Verifier applied strict penalties to the yield estimates. These penalties were reductions in estimated yield based on likely farmer noncompliance to competitor packages. It may also be because of differences in our comparison group. The Verifier used average practices according to a baseline survey instead of collecting data on concurrent average practices in non-AgResults areas. Also, the Verifier used information from baseline farmers' applying similar practices (for example applying fertiliser the same number of times that the AgResults competitor's technology package required) rather than using the information from the baseline to construct a representative average practice.

In contrast, we constructed a comparison group by directly speaking to farmers to ask them about their practices in that season; we selected similar farmers on broad-brush household-level and plot-level characteristics such as land size, flatness, and use of machines. There is always the possibility of unobserved characteristics between the AgResults farmers and the comparison farmers that the evaluation cannot take into account, but we believe a more crucial difference between our comparison group and that of the Verifier is that our comparison is a sample of what occurred 'on the ground' in 2020. It does not only include farmers that use the same rice varieties as those in AgResults, and observed drainage schedules were also different from those found in the Verifier's baseline survey according to our conversation with the Verifier.

Although they do not align with the Verifier's 2020 estimates, our results align with the Verifier's finding that the technology packages had increased yield by 8% and 23% in the spring and summer of 2019. Farmer attitudes about technology packages, discussed in Section 4.3, align with the strongly suggestive evidence that the competitors' technology packages increase yield. We conclude that the evidence strongly suggests that the use of competitors' technology packages and their AgResults-related technical support activities increased yield in the 2020 crops.

5.4 Gender differences in revenue, yield, and decision-making

Based on responses to the survey in the matched comparison study, we did not find gendered patterns to revenue, yield, or decision-making outcomes. Exhibit E-5 in Annex E displays the differential impact of technology package uptake for female- and male-headed households.

Revenue: We examined whether revenue impacts were different for male-headed households than female-headed households and did not find any consistent patterns. There is no statistically significant difference in the percentage of female-headed households that sell their rice. While there are small, statistically significant differences for sales volume, price obtained, and net value, they are not consistently in one direction or the other, and they are not consistent across seasons.

Yield: Working with competitors led to a larger increase in rice yield for female-headed households than for male-headed households. However, there are no clear gender-differentiated patterns of technology package uptake to explain this difference. Working with competitors led to a larger decrease in the number of times fertiliser was applied for female-headed households than for male-headed households. However, this difference alone does not explain why working with competitors led to a larger yield increase for female-headed households than for male-headed households.

Decision-making: There were no gender-differentiated changes to decision-making due to AgResults, measured through questions on who makes decisions on what to do with rice or income from rice in our spring and summer farmer surveys.

5.5 Changes in emissions

Our reviewed identified anomalies in the Verifier's emissions data, implying that it is not possible to ascertain the extent to which emissions may have increased or decreased. The difficulties in estimating emissions offer important lessons that will help future projects. This project was a uniquely large-scale effort to measure agricultural emissions from smallholder farmer plots. As far as we are aware, there are no similar large-scale studies of emissions reduction for alternate wetting and drying, which was a key component of the AgResults competitors' technology packages. GHG emissions are difficult to model and predict, particularly in developing countries where it can be difficult to find trained technicians and where measurements are collected from small plots near other small plots using different technologies. We first present evidence that the estimates are unreliable. We next offer recommendations for approaches to estimating emissions in future technology packages. Finally, we present and describe emissions reductions as estimated by the Verifier.

We believe the estimates are unreliable for several reasons:

- The Phase 1 interim estimates were reported by the Verifier as unreliable—so unreliable it was decided not to use them in determining Phase 1 interim prizes after Phase 1 Crop 1.
- When Phase 1 estimates were used to determine Phase 1 Crop 2's emissions, they were still uncertain. Even for competitors whose estimated emissions reductions were high on average, the uncertainty was so high that some of those estimates were not statistically significantly different from zero—it was possible the high estimates were due to chance (Abt Associates, 2019).

- Despite the uncertainty of the Phase 1 Crop 1 estimates and the limited number of plots used in Phase 1, the Verifier relied on measurements from Phase 1 Crop 1 and Phase 1 Crop 2 to refine the model that was used moving into Phase 2. The refined model was never re-tested in field conditions using the types of measurement inputs used Phase 2.
- We hired the Verifier to use the data collected in the farmer diaries to estimate average emissions in Thai Binh province, but the results were not within a reasonable range of emissions reductions. Emissions estimates using the diary data from AgResults farmers were roughly double the findings that the Verifier had estimated for AgResults farmers for the project's verification, even though these two sets of emissions estimates should have been similar.[17] (These high emissions findings may be due to differences in how water use was reported. That the model cannot handle average farmer water usage reports to provide estimates is a sign of potential problems).[18]
- There are no correlations within the diary study between emissions reductions modelled by the Verifier and practices highly correlated with emissions reductions. For example, water usage should account for a high proportion of emissions

Summary of Verifier's emissions measurement methodology

The Verifier used a mixture of direct measurement and modelled estimates to gauge emissions. In Phase 1, the Verifier based emissions estimates on modelled and directly measured results from 22 test plots (a control plot and a treatment plot for each of the 11 Phase 1 competitors), using an internationally accepted Denitrification Decomposition (DNDC) model. The comparison plots used the same rice varieties as the competitors but counterfactual practices. These counterfactual practices were based on average practices in Thai Binh according to a survey that Applied Geosolutions oversaw of 720 households, randomised by production area and soil type (Salas, 2017, 2018). Phase 2 estimates used the Phase 1 counterfactuals, and the treatment, inputs (such as drain times and planting dates) provided by cooperative leaders as well as random site checks of farmer planting and fertiliser practices. The firm assigned compliance scores to competitors for drainage and planting practices, and incorporated these scores into its modelled emissions.

---

[17] The Verifier reported average emissions among AgResults farmers of 3.79 MT/ha in the spring of 2020 and stated that these levels were a 3.95% reduction in average emissions, implying the average farmer in Thai Binh had emissions of 3.9 MT/ha. For summer 2019, the Verifier reported average emissions among AgResults farmers of 6.79 MT/ha and stated that these levels were a 3.35% reduction in average emissions, implying the average farmer in Thai Binh had emissions of 7.0 MT/ha. Therefore, the commissioned estimates we received with average emissions twice that of AgResults farmers (roughly 7 MT/ha in spring and 14 MT/ha in summer) are quite problematic.

[18] The discrepancies are likely a combination of (a) a difference between the Verifier's assumptions about the farmers' cultivation practices absent AgResults technology and the actual practices of the farmers who participated in the diary study and (b) the model's sensitivity to input data. Information in the diary differed from the kind of information collected by the Verifier to estimate emissions from the AgResults farmers. For example, the Verifier provided as input to the model all of the scheduled drain events of the AgResults farmers, as well as measures of uncertainty as to whether the drain events truly occurred. Farmers completing diaries recorded every time water levels dropped to zero naturally, even though the fields were almost always immediately flooded again. Although these are 'drain' events, most of these events were natural rather than forced drain events of already-flooded fields. Thus, for the diary study the model considered frequently fluctuating water levels (on average 11 or more drain events), whereas for the AgResults study the model's inputs about water management were limited to the number of forced drain events of already-flooded fields (usually between three and six drain events).

reductions according to the Verifier, yet reduced water usage based on phase of the crop cycle in the diary does not correlate with reduced emissions.[19]

- The Project Manager reported that at the end of Phase 2 the "Verification system had [a] high level of uncertainty and was not internationally accredited i.e., by gold Standard/Verra Standards to trade in the carbon market as additional revenue for competitors" (Steering Committee, 2021).
- The uncertainty in the findings complements the reported confusion of cooperative leaders, farmers, and even competitors about basic best practices to lower GHG emissions.

The Verifier followed a number of best practices, including using real-world field demonstrations (not in laboratory) in Phase 1 and basing comparison plots on the results of a randomised survey of farmer practices. To build on those practices in the future when developing new technologies, we recommend:

- The innovation (Phase 1) and dissemination (Phase 2) phases would ideally use the same means of measurement to more clearly identify whether differences across phases are due to measurement or actual changes. If this is not possible, testing should determine how results change due measurement differences.
- Phase 1 testing should occur with sufficient sample size to avoid over-reliance on one or two demonstration plots' results. Power calculations can determine the sample size necessary to result in reliable estimates.
- If model results are confusing, the model should be tested again on new data in similar conditions (such as the same season).

With these caveats in mind, we recap the emissions estimates as reported by the Verifier. The business plan's "conservative estimate of emissions reductions" anticipated a 30% reduction in GHG emissions. Phase 1 estimates for reduction ranged from 40% to 100% for the competitors that proceeded to Phase 2 (Exhibit 5-4).

Exhibit 5-4. GHG performance of Phase 1 technologies (as reported by the Verifier)

Competitor	Total GHG emissions reduction (MT CO_2e/ha, sum of spring and summer emissions)	Total GHG emissions reduction as percentage of average in Thai Binh (set at 10.66 MT CO_2e/ha)
I4	7.9	74%
I5	10.8	101%
I18	4.4	41%
I23	6.5	61%

Source: Abt Associates (2019).

[19] The correlation between water use and Verifier GHG emissions in the income study data (for a small sample size of cooperatives that were in both the income survey and the Verifier estimates) is also not strong or clear. Using a linear regression, we found no link between the percentage of the crop cycle during which the field was dry and the estimated GHG emissions in the diary data. We also found that lower planting density was significantly predictive of lower GHG emissions but accounted for only 2% of the variation in GHG emissions. We found no link between GHG emissions and fertiliser application (we jointly tested number of times applied, nitrogen application rate, and whether application adhered to AgResults recommendations). We make no claim that GHG emissions are independent of these factors, only that it is surprising that the link is not more obvious in a simple model. The lack of a strong correlation could be due to errors in the data, including the possibility of errors in the GHG estimations.

The Verifier reported that both emissions reduction per farmer and total emissions reductions in Phase 2 were substantially less than anticipated based on Phase 1 emissions findings. Exhibit 5-5 displays the emissions reduction potential estimated in Phase 1 and compares it to the realised emissions reductions for each of the four crops in Phase 2. In the spring crops, only one implementer (i4) ever reached or exceeded its estimated mitigation potential. In the summer crops, no implementer ever reached or exceeded its estimated mitigation potential: the closest is i18 in Crop 2, which achieved 35% of its emission reduction potential (0.82/2.60). The range of performance is quite wide when broke down by season and competitor.

Exhibit 5-5. Competitor technology packages' emissions reduction in Phase 2 compared to Phase 1

Emissions reduction (MT CO_2e per hectare)	I4	I5	I18	I23
Phase 1 Spring, estimated mitigation potential	**0.82**	**1.34**	**1.76**	**0.67**
Phase 2 Crop 1 (Spring)	1.84	0.35	0.77	-0.03
Phase 2 Crop 3 (Spring)	0.20	0.19	-0.14	0.23
Phase 1 Summer, estimated mitigation potential	**7.07**	**9.45**	**2.60**	**5.78**
Phase 2 Crop 2 (Summer)	1.92	1.32	0.82	1.11
Phase 2 Crop 4 (Summer)	1.00	0.55	-0.53	-0.36

Source: Secretariat reports prepared for the Steering Committee Meeting, October 2018 and Spring 2021.

As reported by the Verifier, the total emissions reductions across the four AgResults seasons were 270, 1225, 203, and 799 MT of CO_2e. Vietnam's average annual emissions per person is 1.93 MT of CO_2e; thus, based on these data the reported reductions would be equivalent to the emissions of 1293 Vietnamese citizens total, or 140, 634, 105, and 414 Vietnamese citizens for the four AgResults seasons, respectively.

The Verifier estimates that without the application of strict compliance penalty scores to account for imperfect use of technology packages, the results would be a reduction of 1.57 MT CO2e/hectare, or about double the reported overall estimated rate of 0.67 MT of CO_2e/hectare in Phase 2 (but less than 25% of the overall estimated rate in Phase 1).

6. Evaluation Question 4: Evidence for sustainability

Two of the four competitors continued to invest in the first season following the competition. Both are very likely to continue to invest in the dissemination of their technology packages given the alignment between the technology packages and their core business models. Farmers were favourable about continuing to use the technology package, although many emphasised that continued use of the packages depended on continuing engagement with the competitors that supported their use of the packages during the project. Government actors at the national, provincial, and local levels were also positive about continued promotion of the technology packages; an important factor given the strong role played by the public sector in Vietnam. There is little evidence to suggest that the project will lead to near-future widespread changes to production and markets.

Sustainability is likely to be driven by competitors and farmers' perceptions of benefits of the technology packages, and the degree to which local governments support continued promotion of the systems.

Competitors reported intent and actions to continue to promote their technology packages. The levels of intent and action were clearly aligned with the degree to which continued use of the technology practices aligned with their core business activities. The technology packages of AgResults competitors i4 and i5 were described by the competitors as closely integrated and consistent with their core business activities—procurement of specialty rice and rice for processing into seed, respectively. Both have stated that the technology packages enhanced their procurement efforts. For example, in an email to AgResults sponsors and other stakeholders, the CEO of i4 stated that participating in the project "…is in line with the goals and activities that I4 (has) implemented in TB province: That is to build a process to produce high quality rice on a large scale, control fertilisers and plant protection drugs to meet export standards to markets such as EU, US, Japan…" (email from i4, 3 June 2021).[20] In contrast, it is difficult to identify an underlying business case for i18 to continue to promote their package. They departed from their plan to promote their own rice varieties and instead promoted the use of a competitor's rice variety to maximize uptake among farmers and qualify for the AgResults prize. Sustained application of i23's technology package may be limited by the relatively high prices of the package's inputs—in our qualitative interviews with farmers, numerous farmers expressed reluctance to pay more for the inputs, despite also expressing appreciation of their benefits.

Farmers appear to be more consistently positive about continuing to apply the technology packages, or at least components of the packages. As discussed in Section 4.4, farmers described numerous benefits of different aspects of the systems, including reduced expenditures, reduced labour requirements, improved quality and yields, and in some cases also having their rice purchased by competitors at favourable prices. They also received training and in-kind credit to facilitate their correct application of the technology packages. Our qualitative endline interviews with farmers working with each competitor revealed that farmers were largely very happy with the technology packages and their components, but when asked about continuing to apply the packages as a whole, they frequently made reference to the need for continued support from the competitor to enable their continued use. Some farmers also described themselves as being unwilling to pay higher prices for inputs such as fertiliser, despite the benefits that they perceived from them. This reluctance suggests that without continued price discounting of the more expensive inputs, individual components of the larger technology package may be discontinued. These qualitative findings align with the quantitative results showing that net value of rice production is not

[20] We understand i4's references to 'large scale' to mean relative to its current operations.

improved without price discounts on average, across seasons. In general, farmers were most interested about using the packages when they received market benefits such as purchase commitments, premium prices, services such as collection of freshly harvested rice paddy from the field (saving them from having to transport, thresh, and dry the rice), and production inputs and credit.

Continued government support, at multiple levels, will also be important to sustaining the technology packages. A number of farmers stated that their sustained use of the technology packages would be influenced by the recommendations of their cooperative leaders or local extension leadership. An important consideration with respect to sustainability (and replicability in other areas) is the organisation and coordination of farmers with contiguous plots—required by AgResults. Cooperative leaders were crucial partners in this organisation, and while some had challenges with reluctant farmers, most cooperative leaders appreciated the resulting economics of scale. As one cooperative leader summarised: “The rice is blossoming at the same time; high uniformity leads to easy harvesting by machine.” Another cooperative leader reported, “The cooperative operates smoothly because farmers transplant in large fields…this reduced the cost of reaper and reduced the rice dropped in the field… [and reduced] the cost of transplanting (reduced rice seed, less workdays due to sparse transplanting).” Cooperative leaders may continue to receive encouragement from provincial and national levels, in addition to the crucial technical and financial assistance provided by competitors. From a higher-level policy standpoint, the GoV and provincial government have a stated commitment to the promotion of emissions-reducing technology packages. Our interviews of Ministry of Agriculture and Rural Development representatives on the Advisory Council also provide evidence for continued support for the promotion of technology packages that maintain or enhance farmers’ welfare while reducing GHG emissions.

Local coordination and leadership to support sustained use of the technology packages is required in part because they are very specific and multi-dimensional. Our interviews revealed that farmers and cooperative leaders require extensive support and deadline reminders to follow the technology package schedules and plans. Competitors and/or scientists involved in developing the technology packages likely need to be involved in explaining their specific package given the differences between them. Even for technology packages involving the same variety, each competitor had different planting density, fertiliser, and water drainage requirements. Thus, farmers will need a lot of directed leadership to prevent confusion and partial uptake.

Another consideration with respect to sustainability (and replicability in other areas) is the ability of farmers and cooperatives to manage water. The drainage systems are public utilities and thus not an area of investment that a private sector actor would pursue. As shown in Section 4, the water management component of the AgResults technology packages is the component most different from standard practice. Forty-five percent of cooperative leaders interviewed stated that water management is a challenge. They cited difficulties persuading farmers to use less water, but also cited difficulties with infrastructure including old drainage systems and non-flat fields. Central authorities’ encouragement of technology package uptake may be best matched by simultaneous investments in infrastructure to support the technology packages’ water management requirements.

7. Evaluation Question 5: Cost-effectiveness

Given the uncertainty of the emissions reduction measurements, we do not report a specific cost per MT of CO_2e reduced. In terms of per-farmer and per-hectare costs to deliver the technology, this PfR project had costs similar to non-PfR projects in Vietnam ($81 per farmer and $747 per hectare).

The total cost of the awards, verification, and in-country management in 2017 U.S. dollars, with discounting, was $3,572,778, with 33% spent on awards, 39% on verification, and 29% on in-country management. The cost-effectiveness analysis excludes design costs and the Secretariat's costs, as these costs were not available to the Evaluator. We analyse the costs in 2017 U.S. dollars, which was the mid-point of the project and also the same year for which the External Evaluator published cost-effectiveness data for the AgResults Nigeria and AgResults Kenya projects. We use a 12% discount rate, in keeping with standard discount rates of FCDO and the U.S. Agency for International Development.

This chapter analyses the per-unit cost of achieved outcomes. Exhibit 7-1 displays the per-unit costs for various outcomes achieved, considering various combinations of Phase 1 and Phase 2 costs.

Exhibit 7-1. Cost per unit of outcome achieved

Unit of Outcome	Phase 1 and Phase 2 costs, excluding design & Secretariat costs	Phase 2 costs, excluding design & Secretariat costs
Cost per farmer per season	$86	$55
Cost per unique farmer who used an AgResults technology package at least once	$127	$81
Cost per hectare	$747	$474

Note: Number of farmers per season, number of unique farmers reached, and hectares under AgResults are estimates provided by the Verifier.
The Verifier also provided estimates of total MT of CO_2e reduced. Section 5.5 explains the uncertainty of those estimates. Due to the uncertainty, we do not place emphasis on the cost per MT of CO_2e reduced, which is estimated at $1432 (inclusive of both Phase 1 and Phase 2 costs) or $910 (inclusive of only Phase 2 costs). For comparison, carbon offset calculators for airplane travellers looking to offset their contribution towards emissions from their flights currently quote offset prices ranging from $10 per MT to $90 per MT, depending on the offset activity. Two examples of carbon offset opportunities are https://carbonfund.org/carbon-offsets/ and https://co2.myclimate.org/en/flight_calculators/new.

Focusing on technology development alone, the cost per Phase 1 technology package that was deemed successful and eligible for Phase 2 was $217,063.[21] The prize opportunity enticed 24 competitors to apply to participate in Phase 1 and resulted in six technologies that were deemed to meet requirements. The prize was sufficiently large and the technology requirements seemingly feasible to solicit high participation and promising technology packages. We do not have comparable costs for the development of other technology packages outside of AgResults.

With regards to technology transfer, prize sponsors effectively paid $474 in project management costs, outcome verification costs, and prize awards for every time one hectare of rice was cultivated using an AgResults technology package in Phase 2 of the project. Farmers' discounted average impact of net value per hectare of rice was $39.[22] Although

[21] This figure includes in-country project management, outcome verification, and prize award costs, and excludes project design costs and Secretariat costs.
[22] The average impact on net value, reported per sao in Section 5.1, is 2361 in '000 VND/hectare. As done for the cost estimates, we convert this figure to 2017 dollars, and then adjust for discounting to the first year of the project.

$474 is much larger than $39, the per-hectare project cost is roughly in line with other emissions reduction initiatives. In 2019, CGIAR proposed a $371M strategic plan to establish AWD practice on 900,000 hectares in the Mekong Delta, which is a cost of $412 per hectare and includes long-lasting infrastructure upgrades (Tran et al., 2019). The World Bank anticipated influencing rice farmers over 75,000 hectares to adopt the One Must Do, Five Reductions model in the Mekong Delta for a cost of $1,495 per hectare, which includes ongoing grant support to newly established farmer organizations.

The cost-effectiveness measures presented in Exhibit 7-1 are limited in that they explain the cost-effectiveness of outcomes to date. They are also limited in that they do not highlight the cost per MT of Co_2e reduced, owing to our uncertainty of the Verifier's estimates. If future technology package uptake and reduced emissions are attributable to AgResults, revisions to the cost-effectiveness measures would improve the favourability of the project.

In the immediate future, costs of any future technology package uptake will likely be incurred entirely by the private sector. We estimated an approximate cost to competitors of $151 per hectare moving forward. This estimate was based on information from one of the competitors, which provided their costs (mostly staff oversight and training) and price discounts offered in 2020 when working with AgResults compared to non-AgResults farmers. Thus, future uptake would likely depend on the competitor's gain of at least $151 per hectare in revenue using the AgResults technology compared to non AgResults technology to cover costs. As two competitors engaged roughly 6000 farmers each in the season after AgResults ended, they must expect additional revenues from the AgResults technologies that covers costs—even absent the AgResults prize money.

Adequate ability to drain is a key requirement of the AgResults technology packages. Although improvement of drainage systems would be a prohibitively expensive undertaking for any of the private sector actors involved in AgResults, it could serve as an alternative donor investment to the prize competition. Vietnam is increasingly susceptible to damaging floods due to climate change, and our data suggest that the drainage infrastructure does not facilitate easy adherence to AWD-type requirements in the monsoon seasons (i.e., the summer crop). One alternative approach to promoting technology package components in Thai Binh could be to sponsor infrastructure improvements. To enable farmers to practice AWD, the World Bank budgeted almost 100 times as much ($182M) as AgResults Phase 1 to improve public infrastructure, notably drainage systems, across an area 20 times the size of Thai Binh province (World Bank Group, 2015). Assuming a constant ratio of land area to infrastructure improvement costs, $9M (in 2017 U.S. dollars) would go a long way towards improving drainage infrastructure in Thai Binh.[23] The total discounted cost of the AgResults Vietnam project for both Phase 1 and Phase 2 was $5.1M.

[23] We obtain this estimate by multiplying the dollar/area ratio (World Bank expenditure divided by area of land mitigated) by the area of Thai Binh province.

8. Evaluation Question 6: Lessons

We provide lessons learnt according to key facets of the prize competition.

Private sector involvement. The AgResults Vietnam Challenge Project demonstrated that PfR approaches can be successful at spurring the private sector to address climate change. In Thai Binh, AgResults competitors increased the supply of GHG-reducing inputs and technology packages, the number of farmers using the technology packages, and the availability of rice produced using these packages.

The prize competition did not provide direct incentives to develop markets for the technology packages or their derived goods (carbon credits, rice valued for attributes that are produced using the promoted technology packages) that would help to sustain the use of technology packages after the end of the prize competition. As a result, private sector commitment to sustaining the use of the newly developed technology packages is mixed. The two best-performing competitors developed these market linkages independently.

Public-private sector collaboration. This project demonstrated that PfR projects can motivate the private sector to affect public sector action. Like many climate initiatives, GHG emissions reduction in northern Vietnam required action and leadership from the public sector. An important driver of GHG emissions reductions, water usage, is not market-based but rather in the hands of local government officials. In some cases, competitors' close collaboration with the public sector, particularly cooperative leaders, led to changes in water management. Farmers and cooperative leaders agreed to reduce water usage in anticipation of increased yields from working with competitors. The increased yield in turn benefited some of the competitors that sold rice products. This suggests that a particular synergy may come from PfR public–private sector collaboration compared to direct investment in public sector works and services. To understand this synergy, it could be fruitful to compare the AgResults project to the World Bank-funded One Must Do, Five Reductions program in Vietnam, once it releases GHG emissions information (Jackson et al., 2015).

Verification. Prize sponsors and competitors should be satisfied that verification procedures are valid and reliable. This is especially important to ensure reliability of estimates of main outcome measures in a development phase before continuing to the dissemination phase. The prize structure placed very strong assumptions on the validity of the verification in Phase 1, resulting in Phase 2 awards being paid out for dissemination success despite the project's likely not achieving its target emissions reductions.

Competitors may also need time during the innovation phase to adjust their innovations after testing. Competitors did not learn whether their technologies reduced emissions until prizes were announced, and no competitor had time or resources to first experiment systematically with alternate approaches.

The verifier estimating GHG emissions should also apply best practices related to sample size and consistency both in types of measurement over time and in testing under similar conditions multiple times (i.e., in the same season).

Innovation. PfR projects can create a diversity of innovations. The objective of the AgResults Vietnam Challenge Project was never to produce a singular 'best practice' that is easy to communicate and promote. However, as a consequence, it allowed for the promotion of packages that differ markedly in guidelines even for a single rice variety in the same season. As the technology packages are complex and multi-faceted, there is a potential for ambiguity among farmers and cooperative leaders about 'best practice', with differing guidelines perhaps implying that variations are acceptable. Future work on emissions reduction practices for rice could refine the technology packages and further study which are the best at reducing emissions.

In addition, the long-term sustainability of these highly technical practices should be compared to the sustainability of other recent GHG-reducing innovations that require less oversight. As a comparison, for example, one could study the use of films laid on fields to help retain water so as to flood fields less often while maintaining or improving yields. Plastic films (which pollute the ground) are already widespread in parts of China and have reduced GHG emissions. Biodegradable films are still being developed for commercial use (Yao et al., 2017).

Prize structure. Future prize competitions should carefully consider how to best incentivise achievement of the main development outcome. The AgResults Vietnam Challenge Project's prize structure, where winners received prizes proportionate to their results, allowed sponsors to promote multiple outcomes—smallholder uptake (40%), repeat use by farmers (20%), yields (20%), and GHG reductions (20%). The grand prize winner ultimately won by disseminating its technology package to large numbers of repeat farmers, but it is unclear whether it substantially reduced emissions. According to the Verifier's emissions estimates, a different competitor had much more substantial emissions reductions and similar yield outcomes, but did not win the prize competition.

The weighting also led to one competitor "gaming the system" to reach more farmers by breaking up plots among family members and registering each family member as a unique farmer. Potential solutions to these challenges involve alternative weighting to emphasise overall success at reducing GHG emissions (e.g., rewards based on total emissions reductions across all farmers, irrespective of the number of farmers engaged), and making sure that all competitors clearly understand the definition of a 'unique farmer.'

Since GHG emissions in Phase 2 were uncertain, future prize competitions should carefully consider how to best incentivise clear GHG emissions reductions. For example, some competitors in Phase 1 did not move on to Phase 2 despite having high estimated success in reducing emissions because of greater variation in yield outcomes: in some seasons yield increased, but in others yield decreased. Emissions reductions may have been more successful if the prize competition had allowed those competitors to operate in the seasons for which they increased yield. In addition, prize sponsors could avoid paying prizes for below-target emissions reductions by incorporating minimum threshold amounts.

Revenue was another overlooked measure of farmer well-being that may have filtered higher-performing competitors from Phase 1 into Phase 2. Some competitors could potentially increase revenue even while not increasing yield through growing higher value and/or lower cost products.

Theory of Change. Emphasising alignment of technology packages with market opportunities could help to increase uptake and sustainability of GHG-reducing technology packages in the future. The AgResults Vietnam Challenge Project's implicit theory of change assumed that there was no financial incentive—beyond the AgResults prize—for competitors or farmers to work with GHG-reducing technology packages. Within the prize competition, AgResults fostered the creation of the competitor-specific market channels that competitors used to supply farmers with technology package inputs and systems, take possession of the rice produced using those systems, and then feed it into their existing market channels. Alternatives to competitor-specific markets could be markets for GHG-reducing inputs such as slow-release fertiliser; markets for rice produced using GHG-reducing technology, and carbon offset markets.

For future prize competitions, implementers could also consider restricting competitors to those with a clear business case for investing in the technology packages. I4 is an example of a company that leveraged the prize competition to produce rice for the specialty markets it served in part because stringent requirements of those markets align with the technology packages' requirements. I4 is also one of the two companies that continued to promote its newly developed technology packages after the AgResults project ended.

9. Conclusion

The AgResults Vietnam prize competition successfully spurred the creation of GHG-reducing technology packages. The extent to which these technology packages reduce emissions is unclear due to difficulties with emissions measurement. The prize competition successfully promoted yield increases of 14%.

The evaluation assessed the project's impact on private sector involvement in the dissemination of emissions-reducing/yield-enhancing technology packages and the adoption of these packages by smallholder farmers. It also assessed subsequent impacts of adoption on other smallholder outcomes, the project's cost-effectiveness, and the sustainability and scale of impact. The evaluation findings are summarised below.

Private sector involvement. AgResults spurred substantial investment by the AgResults competitors to scale up dissemination of their technology packages. Each competitor contracted either cooperatives or individual farmers to produce rice using their respective systems, and supported that production with training, services, and reduced input prices. Two of the four Phase 2 competitors incorporated their AgResults-developed technology packages into their core rice business activities.

Adoption. The scale of adoption is similar in scale to other comparable projects (although much lower than the target of 75,000), as measured by the adoption rate of 28,031 unique farmers. Given AgResults competitors' limited reach relative to the roughly 483,000 rice farmers in Thai Binh, AgResults had limited impact on average rice cultivation practices across Thai Binh province as a whole.

In comparing AgResults farmers' practices to comparison farmers', we found that AgResults farmers significantly and substantially changed their practices. The components of the packages that were substantially different from standard practices were the drain schedules/reduced water use, reduced planting density, less burning of fields, less nitrogen use, and more frequent fertiliser applications. Given the largely female-dominated field of rice farming in Vietnam, most adopters were women.

Smallholder yield and income. Compared to the rice plots of the matched comparison group, we found that the plots where AgResults competitors' technology packages were applied had 14% higher yields and 11% higher net value. This result is comparable to findings from one of the only rigorously evaluated rice intensification projects (Barrett et al., 2021).

GHG emission reductions. We find emissions estimates uncertain. According to the Verifier's estimates, one technology package of the four used provided substantial reductions in GHG emissions (approximately 20%) although still lower than the original goal of 30%. The Verifier modelled between 3.4% and 10.4% emissions reductions overall, across four crop seasons. This reported emissions reduction is equivalent to the average annual emissions of 1293 Vietnamese citizens.

Cost-effectiveness. Given the uncertainty of the emissions reduction measurements, we do not report a specific cost per MT of CO_2e reduced. In terms of per-farmer and per-hectare costs to deliver the technology, this PfR project had costs similar to non-PfR projects in Vietnam ($81 per farmer and $747 per hectare).

Sustainability. Project sustainability will rely on the degree to which competitors and farmers feel the technology packages benefit them, and the degree to which local governments (particularly cooperative leaders) provide support. Farmers benefited from the technology package, but likely need reduced input prices to continue; the yield increases are not clearly substantial enough to justify the increased costs for them otherwise. Reduced input prices could possibly be provided indirectly by carbon markets or by the competitors in the future. For competitors, reducing input costs for farmers may be worthwhile depending on their business model and their expected revenue. Cooperatives and other government officials are likely to approve of attempts to continue given the alignment with the “Rice-Restructuring Plan and Crop Production Strategy for 2020-2025 with vision to 2030” and Vietnam’s Nationally Determined Contributions. The technology packages are intended for use on flat land that is capable of draining.

In summary, the project achieved substantial gains in all areas except GHG emissions reductions, where the results are uncertain. The project demonstrated that PfR projects can spur substantial private investment in the development and dissemination of new agricultural technologies. This project is one of the first agricultural emissions reductions projects conducted with large numbers of smallholder farmers. It is a crucial step in generating greater knowledge of emissions reduction measurement in the real world, and the findings should inform future projects. It is clear that improvements in GHG emissions measurement are necessary, especially measurements using indirect data. Future work done on larger scales may need to revise expectations to take into account measurement difficulties and difficulties with real-world applications of emissions-reducing/yield-enhancing technologies.

10. References

Abt Associates 2017. AgResults Vietnam Design Report.

Abt Associates. 2019. AgResults Vietnam Emissions Reduction Project: Phase 1 endline assessment.

Arrow, K. 1970. Political and economic evaluation of social effects and externalities. *The Analysis of Public Output.* pp. 1–30.

Barrett, C., A. Islam, A. Malek, D. Pakrashi, U. Ruthbah. 4 July 2021. Experimental evidence on adoption and impact of the system of rice intensification. *American Journal of Agricultural Economics.* pp. 1-23. https://onlinelibrary.wiley.com/doi/10.1111/ajae.12245

Beckerman, W. 1972. Economist, scientists and environmental catastrophe. *Oxford Economic Papers* 24 (3): 327-344. https://doi.org/10.1093/oxfordjournals.oep.a041225

BRIA II. 2021. Factsheet: Better Rice Initiative Asia (BRIA II) Improving the livelihoods of smallholder rice farmers in Thailand, Vietnam and Indonesia through Sustainable Market Oriented Smallholder Value Chains. https://www.asean-agrifood.org/factsheet-bria2/. Accessed 3 November 2021.

Deloitte. 2017. Secretariat Steering Committee Presentation. March.

Demont, M., and Rutsaert, P. 2017. Restructuring the Vietnamese rice sector: Towards increasing sustainability. *Sustainability* 9, 325*.* February 22. https://www.mdpi.com/2071-1050/9/2/325/pdf

Environmental Defense Fund (EDF). 2018. Global risk assessment of high nitrous oxide emissions from rice production. White Paper. https://www.edf.org/sites/default/files/documents/EDF_White_Paper_Global_Risk_Analysis.pdf

FAOSTAT. 2017. Viet Nam Country Indicators. FAO. Viewed on 1 October 2017. http://www.fao.org/faostat/en/#country/237

Geyer, J., A. Greif, and C. Price. 2020. AgResults Vietnam Emissions Reduction Challenge Project: Phase 2 Randomised Controlled Trial Adherence Report. Abt Associates. Prepared for the U.K. Department for International Development and AgResults Steering Committee.

Guan, K., Hien, N.T., Li, Z., Rao, L.N. 2018. Measuring rice yield from space: The case of Thai Binh province, Viet Nam. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3188560

Jackson, C., C.R. Reddy, and B.T. Cao. 2015. A slogan for sustainable agriculture: 'Mot Phai, Nam Giam' rice production. World Bank Blogs. https://blogs.worldbank.org/voices/slogan-sustainable-agriculture-mot-phai-nam-giam-rice-production-vietnam.

Jain, N., Dubey, R., Dubey, D. S., Singh, J., Khanna, M., Pathak, H., Bhatia, A. 2013. Mitigation of greenhouse gas emissions with system of rice intensification in the Indo-Gangetic plain. *Paddy Water Environ.* doi:10.1007s/10333-013-0390-2.

Khosa, M. K., Sidhu, B. S., Benbi, D. K. 2010. Effect of organic materials and rice cultivars on methane emission from rice field. *Journal of Environmental Biology,* 31:3, pp. 281-285.

Liu, G., Zotarelli, L., Li, Y., Dinkins, D., Wang, Q., Ozores-Hampton, M. 2014. Controlled-release and slow-release fertilizers as nutrient management tools. HS1255. Gainesville, FL: Horticultural Sciences Department, University of Florida/Institute of Food and Agricultural Sciences Extension. http://edis.ifas.ufl.edu/pdffiles/HS/HS125500.pdf

Narayan, Tulika, Judy Geyer, Denise Mainville, Adi Greif, and Cristofer Price. 2020. AgResults Evaluation Design: Vietnam Emissions Reduction Project. Rockville, Maryland: Abt Associates. Prepared for U.K. Department for International Development and AgResults Steering Committee.

Pigou, A. C. 1920. The economics of welfare. London: Macmillan.

Salas, W. AgResults Learning Brief #1. September 2017. The GHG Emissions Reduction Project in Vietnam: Challenges in Defining a Verification Baseline. https://agresults.org/learning

Salas, W. AgResults Learning Brief #2. January 2018. Results and Observations from the Vietnam Challenge Project's First Cropping Season. https://agresults.org/learning

SRI-Rice 2015. International Network and Resource Center. http://sri.ciifad.cornell.edu/countries/vietnam/index.html. Accessed 17 August 2021.

Steering Committee. April 2021. AgResults Steering Committee meeting slides.

Suri, T. 2011. Selection and comparative advantage in technology adoption. *Econometrica* 79(1) p.159-209.

Tivet, F., Boulakia, S. 2017. Climate smart rice cropping systems in Vietnam. State of knowledge and prospects. Montpellier, France: CIRAD, 41 p.

Tran, V.T., V.T. Mai, T.D.T. Nguyen, H.A. Le, M.B. Richards, L. Sebastian, E. Wollenberg, D.Q. Vu, B.O. Sander. 2019. An investment plan for low-emission rice production in the Mekong River Delta region in support of Vietnam's Nationally Determined Contribution to the Paris Agreement. Working Paper No. 263, CGIAR Research Program on Climate Change, Agriculture, and Food Security. https://cgspace.cgiar.org/handle/10568/101315

Truc, N. T. T., Sumalde, Z. M., Espaldon, M. V. O., Pacardo, E. P., Rapera, C. L., & Palis, F. G. 2012. Farmers' awareness and factors affecting adoption of rapid composting in Mekong Delta, Vietnam and Central Luzon, Philippines. *Journal of Environmental Science and Management*, 15(2).

World Bank Group. 2015. International Development Association Project appraisal document on a proposed credit to the Socialist Republic of Vietnam for a sustainable agriculture transformation project. Report No. PAD1168. https://documents1.worldbank.org/curated/en/783441467998463415/pdf/PAD1168-PAD-P145055-IDA-R2015-0175-1-Box391456B-OUO-9.pdf

World Bank Group. 2016. IFC partners with Loc Troi Group to promote sustainable rice production in Vietnam. https://pressroom.ifc.org/all/pages/PressDetail.aspx?ID=24867. Accessed 3 November 2021.

Yao, Z., Zhou, Z., Zheng, X., Xie, B., Mei, B., Wang, R., & Zhu, J. 2010. Effects of organic matter incorporation on nitrous oxide emissions from rice-wheat rotation ecosystems in China. *Plant and Soil,* 327(1-2), 315-330.

Yao, Z., Zheng X., Liu, C. Lin, S. Zuo, Q., Butterback-Bahl, K. 2017. Improving rice production sustainability by reducing water demand and greenhouse gas emissions with biodegradable films. *Sci Rep,* 7, 39855 (2017). https://doi.org/10.1038/srep39855.

Yin, R. K. 2003. Case study research: Design and methods. Thousand Oaks, Calif.: Sage Publications.

Zou, J., Huang, Y., Jiang, J., Zheng, X., & Sass, R. L. 2005. A 3-year field measurement of methane and nitrous oxide emissions from rice paddies in China: Effects of water regime, crop residue, and fertilizer application. *Global Biogeochemical Cycles*, 19(2).

Annex A – Details of study design, regression specification, and sampling weights

Section 2 described the design of the RCT and provided an overview of the supplemental matched comparison study. In this annex we describe:

- The matching technique used for the matched comparison study
- The regression models used for both the RCT and the matched comparison study
- The covariates used in the regression model
- The weighting scheme
- Our process for analysing qualitative data.

A.1 Matching technique

We supplemented the RCT with a matched comparison design that estimate the income impact on farmers of working with an AgResults competitor. As AgResults operated in only a very small area and only with a very small proportion of rice farmers, we chose this approach to better understand the experience of farmers using the AgResults competitors' technology packages. To do so, we surveyed over 1000 farmers listed as working with an AgResults competitor. Their responses are interesting in and of themselves, but to estimate the results of AgResults on their income, we needed to understand what their responses would have been in the absence of the AgResults project. To achieve this, we surveyed over 1000 matched comparison farmers. The matched comparison study's findings are valid for and representative of 85% of the communes in Thai Binh, excluding the 15% of communes that either have little rice-farming area or an extremely high concentration of rice production area.

The selected treatment communes have, on average, larger areas under rice production (in spring and summer) and slightly higher yields. We found that, taken together, these characteristics distinguish the selected treatment communes from the non-selected communes and their differences are statistically significant (the global F test yields a p-value of 0.032). We excluded a few control communes with the lowest areas under rice production at baseline, and a few selected treatment communes with the highest areas under rice production to make them more similar. This design has the 'cost' of having a sample that is representative of roughly 85% of the treatment area instead of 100%.

We used a matched comparison design to assess the impact of technology package uptake on the smallholder outcomes. First, we selected treatment and control cooperatives based on balanced baseline characteristics. Within the treatment communes, we randomly selected farmers that participated in the treatment. We matched comparison farmers to the treatment farmers by selecting a stratified sample that mimicked the competitor selection process, in addition to weighting farmers based on individual characteristics. For three of the four competitors, we chose comparison farmers from control communes. For one of the competitors however, we selected comparison farmers in communes assigned to treatment that did not take up the treatment because that competitor's selection criteria were very selective, and we could not find a sufficient number of farmers meeting those criteria in the control communes.[24] The challenge of a design relying on ex-post identification of the treatment and comparison group is selection bias. There may be underlying reasons why some eligible communes were targeted by AgResults, and those reasons might explain observed differences between the treatment and comparison group even in the absence of AgResults.

---

[24] Competitor i5, at the time of farmer recruitment for Crop 3, only recruited farmers with whom they had previously worked.

Although it is impossible to eliminate the threat of unobserved selection bias, we mitigate selection bias at the farmer and commune level through an explicit understanding of how competitors selected cooperatives with which to work. Competitor selection occurred at the cooperative level and was exogenous to individual farmer characteristics. Typically, competitors asked cooperative leaders to identify a favourable village or neighbourhood with contiguous rice plots. They asked cooperative leaders to coordinate with the owners of the contiguous plots to create a site of at least 1500 square meters. Depending on the competitor, the site also had to meet additional specifications such as having a high elevation/flat terrain or otherwise very good water drainage control, being situated near a main access road, and/or including farmers already known to the competitor. Based on our interviews with competitors, we understood that competitors considered large contiguous areas crucial to the project for ease of draining and for economies of scale, since the competitors might provide inputs, training, monitoring, and might also buy rice directly from the farmer area at the end of the season. The sample recruitment plan was based on this information and is described in Annex B.

Competitor i5 provides reassuring evidence against commune-level bias for the 'comparison' communes taken from the communes assigned to treatment. This competitor initially only worked with farmers with whom they had previously worked; these were only in the communes assigned to treatment. Company technicians assure us that no underlying differences relevant to the outcomes in communes caused them to reach out to some communes assigned to treatment rather than others. Rather, they had reached out to all communes with which they worked that planted the rice varieties included in one of i5's technology packages. As for the rest of the communes, some of them are switching rice varieties so as to continue the program. Others ended up working with another competitor, showcasing their suitability for program inclusion.

We used observable characteristics to mitigate commune-level bias. The selected treatment communes have, on average, larger areas under rice production (in spring and summer) and slightly higher yields. We found that, taken together, these characteristics distinguish the selected treatment communes from the non-selected communes and their differences are statistically significant (the global F test yields a p-value of 0.032). We excluded a few control communes with the lowest areas under rice production at baseline, and a few selected treatment communes with the highest areas under rice production to make them more similar. This design has the 'cost' of having a sample that is representative of roughly 85% of the treatment area instead of 100%. To mitigate concern that the treatment sample may not be fully representative of the competitors' engagement with AgResults farmers, we stratified the treatment and comparison sample by competitor such that the number of communes served by each competitor in the sample will be in the same proportions as the number of communes served by each competitor in the full population of communes served by the AgResults project. Annex C presents the full details of the balance test of the treatment and comparison group.

Using data from farmer surveys (described below), we compared farmers verified as working with a competitor to farmers in the matched comparison group ('comparison group') by examining differences in farmers' average expenditures, revenues, and farming practices. We used regression analysis to control for baseline characteristics, to weight the treatment group so that it is representative of all farmers who used competitors' technology packages, to weight the comparison group so that it is more similar to the weighted treatment group, and to handle inter-cluster correlation at the village level.

A.2 Regression specification

RCT study of technology adoption

We used a linear regression model to estimate the impact of assignment to the treatment group. Assignment is at the commune level; data are at the farmer level. We used commune-level baseline covariates (detailed in Section A.2). The model estimated cluster-robust standard errors with respect to clustering at the commune level. Sampling was stratified by commune, and within communes with AgResults farmers we noticed that AgResults farmers appeared to be over-represented. To generate estimates that are representative of the average farmer eligible to participate in the diary study, we used farmer-level weights. Section A.3 describes the process of selecting the weights.

Matched comparison study of adoption's impact on income

Similar to the RCT study of technology adoption, we used a linear regression model to study all outcomes. Recruitment and coordination of farmers into 'teams' with contiguous land, and thus occurred at the village level. Therefore, the model estimated cluster-robust standard errors with respect to clustering at the village level. Higher levels of clustering were not optimal because some cooperatives had multiple villages working with different AgResults competitors, and correlated outcomes would arise from farmers using coordinated practices. We also used covariates and weights, as specified in Sections A.2 and A.3, respectively.

A.3 Covariate selection

RCT study of technology adoption

The covariates in our impact regressions control for cooperative-level and farmer-level characteristics. We selected the following baseline cooperative-level covariates because the difference between their averages in the treatment and control communes was significantly different when comparing the 50 control communes to the 250 treatment comments, even though the differences are not significant when examining data at the farmer level (Annex C) and using the regression analysis weights:

- Average rice yield in 2018
- Percent of rice farmers who were members in the cooperative in 2018
- Percent of farmers with at least 1500 square meters for cultivating rice in 2018
- Average rice area of a farmer in that cooperative in 2018
- Percent of the cooperative area used for rice cultivation in 2018
- Whether the cooperative owned a riding transplanter in 2018

We also included the following cooperative-level covariates that explain weather, carbon content of the soil, and travel time to an urban centre:

Average precipitation
Average minimum temperature
Soil organic carbon content at 5-15 cm depth
Evapotranspiration
Travel time to a city of 50K-100,000 persons

We also included the following farmer-level covariates, which we do not believe to be endogenous, or impacted by AgResults. Note that farmers did not have choice over which plots to include as coop leaders/competitors demarcated land use, so we do not consider plot-level controls endogenous. There are also too few respondents in AgResults for any plot level selection to make a difference to results (results are the same with exclusion of plot-level controls).

Age
Sex
Total paddy land worked
Rice plot is at low/medium/high elevation
Farmer-reported drainage quality
Farmer-reported soil type

Matched comparison study of adoption's impact on income

The main explanatory variable of interest is the indicator variable for whether a farmer was verified as working with an AgResults competitor to use an AgResults-approved technology package. We used covariates to control for cooperative-level and farmer-level characteristics. For the cooperative-level variables, we included covariates used in the diary survey study as well as the following characteristics on which the treatment and comparison groups would be imbalanced (difference measured in standard effect size > 0.25) without the use of analysis weights:

- Whether one of the top three most common seeds grown in the cooperative in 2018 (pre-AgResults) was an AgResults seed
- Whether the cooperative had a completed irrigation system in 2018
- Average value of rice crop per farmer in 2018

We used the following farmer-level covariates which we do not believe to be endogenous, or impacted by AgResults:

- Whether rice is a main source of income
- Whether the household completed a secondary education
- Number of assets owned
- Number of crop types grown
- Number of animal types owned
- Total area of rice paddies owned

A.4 Analysis weights

RCT study of technology adoption

We use analysis weights in the analysis of the diary data so that the statistics we present are representative of rice farmers in Thai Binh similar to those who responded to the diary: who live in the province throughout the rice season, sell their rice harvest for income, and who have completed a least a primary education and are able to understand and follow the instructions for keeping a diary about their farming practices. The sample without weights is not representative because the sample was stratified: we sampled two farmers in every cooperative that we reached, regardless of the number of rice farmers in that commune. Thus, cooperatives with a larger number of farmers are underrepresented and cooperatives with few farmers are over-represented. In addition, AgResults farmers were more likely to be selected within the AgResults communes, so they are oversampled.

The analysis weight for each observation is equal to the number of farmers for whom that farmer is representative. To calculate the weights, we had to estimate (1) the number of verified AgResults farmers in the commune who are eligible to participate in the diary study and (2) the number of non-AgResults farmers in the cooperative who are eligible to participate in the diary study. The number of verified AgResults farmers is known from the Verification data. Based on the income survey data, we estimate that 10.3% of AgResults farmers met the eligibility criteria to be included in the diary survey. The number of non-AgResults farmers is estimated based on the number of rice farmers in the cooperative as of Spring 2018 minus the number of verified AgResults farmers. Based on our diary survey recruitment effort, we know that 18.5% of rice farmers in the commune meet the eligibility criteria. Exhibit A-1 displays the number of farmers we sampled in the communes randomised to the treatment and control group, and whether or not those farmers were participating in AgResults.

Exhibit A-1. Number of farmers in diary study

Farmer affiliation	Spring 2020		Summer 2020	
	Treatment	Control	Treatment	Control
Not participating in AgResults	349	97	341	98
Participating in AgResults	38	0	46	0

We did not reweight the data to achieve balance on baseline cooperative characteristics because assignment to the treatment and control groups was random. Thus, any differences between treatment and control communes are due to chance. Balance between the treatment and control communes was reported in our RCT Adherence Report, also presented in Annex C of this report (Geyer et al., 2020).

Matched comparison study of adoption's impact on income

Analysis weights in the income survey achieve two goals: (1) they make the treatment group farmers representative of all farmers participating in AgResults, (2) they improve the comparability of the comparison group to the treatment group on baseline characteristics.

$$wgh_{combined} = wgh_{stratification} * wgh_{balance}$$

The stratification weight ($wgh_{stratification}$) makes the treatment group representative of all farmers participating in AgResults. The sample of treatment group respondents is stratified by Competitor, with sample counts displayed in Exhibit A-2. The stratification weight for each sampled AgResults farmer is equal to the number of verified farmers who worked with that farmer's Competitor divided by the number of interviewed farmers who worked with that farmer's competitor. The stratification weight for each sampled *non*-AgResults farmer is equal to the number of all verified AgResults farmers divided by the number of all interviewed *non*-AgResults farmers.

Exhibit A-2. Weighted and unweighted counts of farmers

Farmer affiliation	Spring 2020		Summer 2020		Combined both seasons	
	N	Percent of all AgResults farmers	N	Percent of all AgResults farmers	N	Percent of all AgResults farmers
	All AgResults farmers					
i4	1814	22%	5562	29%	6530	31%
i5	3895	47%	5052	27%	5218	25%
i18	1334	15%	5878	13%	6417	29%
i23	1202	16%	2386	31%	2429	14%
Non-AgResults	Not applicable					
	Total surveyed farmers (income survey)					
i4	300	32%	331	34%	631	35%
i5	272	29%	215	22%	387	21%
i18	200	21%	271	28%	471	26%
i23	173	18%	163	17%	336	18%
Non-AgResults	1236	131%	1090	111%	2326	127%
	Combined weighted count					
i4	1814	22%	5586	30%	6613	32%
i5	3895	47%	5052	27%	5226	25%

Farmer affiliation	Spring 2020		Summer 2020		Combined both seasons	
	N	Percent of all AgResults farmers	N	Percent of all AgResults farmers	N	Percent of all AgResults farmers
i18	1334	16%	5878	31%	2533	31%
i23	1202	15%	2386	13%	6418	12%
Non-AgResults	8244	100%	18,902	100%	20,670	101%

To improve the balance on observable characteristics, we use propensity scores. Prior to balancing, we compared the treatment and comparison groups on all characteristics listed in Section A.3. Of these, the difference between groups exceeded 0.25 standard deviations on the following two measures: whether any of the seed types used in the AgResults technologies was one of that cooperative's top three most common seed types grown in 2018 (prior to AgResults) and whether the irrigation system was complete as of 2018. When we applied the propensity score balancing method to only those two variables, additional imbalanced arose: whether the commune's irrigation system was managed solely by the cooperative leader, average rice yield, and average value of the cooperative's rice crop. We found that a propensity score model with all five of these characteristics achieved reasonable balance. Using propensity scores *p,* we assign a balancing weight ($wgh_{balance}$) equal to *p/(1-p)* for comparison farmers and 1 for treatment farmers. We multiply this weight times the stratification weight to get the final analysis weight for the farmer. Annex C displays the balance tables, which were computed using the weights described here.

A.5 Analysis of qualitative data

Following data collection, transcription, translation, and cleaning, we organised, coded, and analysed the qualitative data. We coded the data around our central analytic themes (thematic coding) and in relation to specific research questions, sub-questions, or analytic objectives (structural coding). The codes were informed by *a priori* concepts drawn from the project theory of change, SCP framework, desk research, and our initial qualitative assessment and baseline results, based on which we development hypotheses about what factors would influence the success of the project. We applied our deductively developed codes to enable content analysis, a form of text analysis that supports qualitative hypothesis testing (Bernard, 2006).

We also analysed data using pattern analysis, in which we evaluated our hypotheses on the basis of field results to ascertain patterns and divergences among similar market actors or segments. The analytic process and interactions with the in-country agricultural economist who either led or participated in all data collection facilitated an active search for disconfirming evidence and alternative explanations for observed outcomes, with further investigation of results that did not align with hypotheses to maximise insight into our results.

We employed best practices to ensure the robustness of our qualitative methods (Yin, 2003 update citation). These included 'naive' questioning approaches (rather than 'leading' questions, which introduce bias), triangulation of data sources (for example, seeking information from multiple levels of the marketing chain to obtain diverse explanations of phenomena), an active search for data that disconfirms our hypotheses, and careful documentation of the evidence supporting results. The validity of the qualitative research is also bolstered by leading with theory-based economic models (such as the SCP framework). These best practices also support nuanced exploration of diverse factors, such as those identified above, that might also affect the project's outcomes of interest.

Annex B – Sampling strategies

B.1 Diary study

To provide a representative sample of farmers across the cooperatives in Thai Binh province, we aimed to recruit two randomly selected farmers per cooperative to record their rice cultivation practices in a structured diary, both in the spring and summer seasons of 2020. Given a context where many farmers are technically listed as plot owners but in practice leave the land fallow or let others work the land, we did not approach farmers directly and instead worked with cooperative leaders to help us find appropriate farmers. To avoid simply selecting friends of the cooperative leader we first randomised the list of farmers in the commune and second, asked the cooperative leader to select the first two farmers that met our criteria.

Our criteria were:

- Farmer can read/write extensively enough to fill out the diary and is willing to do it
- Farmer is knowledgeable about the rice growing practices on the plot and stays in the village to tend to the rice themselves.

Cooperative leaders also had to write down the reasons why they skipped over some farmers. The above criteria were used in the large majority of cases. However, when it was not possible to follow the above criteria, the cooperative leader instead selected the first six farmers who:

- Grew rice via transplanting/livelihood depends on rice (as a sign of being more 'serious' about rice/more likely to be able to track their rice growing practices)
- Sold rice after harvest
- Have a telephone number (landline or mobile)

Out of the six farmers to whom the cooperative leader introduced us, we randomly recruited two, ensuring that their fields were serviced by different drains. In the case where the cooperative leader could not select anyone from the list of 20 farmers provided, they instead introduced us to three farmers.

This process resulted in reaching 233 out of 255 eligible communes for the diary in the spring, despite the interruption of COVID-19. The 22 missing diaries were due to inability to access the commune due to COVID-19 pandemic-related quarantines of certain communes or lack of assistance from cooperative leaders, either because cooperative leaders were busy or did not want to participate, or were new and not allowed by their senior authority to participate, or handovers from old to new coop leaders had not yet been completed. In the summer, we reached 230 communes.

The result of our selection process was not a completely random sample, and instead tended toward farmers with higher levels of education and a focus on rice farming as a livelihood.

B.2 Income survey

We stratified the AgResults sample to obtain a large enough sample working with each competitor to obtain enough statistical power to confidently examine descriptive statistics by competitor. By competitor stratum, we randomly sampled villages where AgResults took place, and within those villages we randomly recruited farmers identified by the Verifier as working with that competitor. We then aimed to select comparison farmers that were similar to the treatment farmers. To do that, we first selected comparison communes that were similar on baseline characteristics to the treatment communes. The selected treatment communes were significantly different from non-selected communes on three out of the 23 variables tested. The selected treatment communes have, on average, larger areas under rice production (in spring and summer) and slightly higher yields. We found that, taken together, these characteristics distinguish the selected treatment communes from the non-selected communes in the treatment group, and their differences are statistically significant (the global F test yields a p-value of 0.032). (See the RCT Adherence Report for details, Geyer et al., 2020).

Exhibit B-1. Commune characteristics by whether competitors worked in them

Baseline commune characteristics	Competitor - selected treatment mean	Non-selected treatment mean	Difference	Signifi-cance (p-value)
Percentage of farmers in commune who belong to a cooperative	79%	87%	-8%	.162
Average farmer rice cultivation area (ha)	0.17	0.16	0.01	.218
Percentage of farmers in commune with >1500 m^2 for rice cultivation	51%	48%	3%	.167
Average rice yield in spring 2018 (MT/hectare)	7.0	6.9	0.2**	.016
Total commune area under rice in spring 2016 crop (ha)	314	275	389**	.006
Total commune area under rice in summer 2017 crop (ha)	315	277	38**	.006
Total harvest per commune (MT)	2207	2023	183	.34
Average value of production per farmer per commune (1,000 VND/MT)	1.2	1.2	0.0	.797
Number of observations	67	138		

Data source: Baseline cooperative-level administrative data.
Notes: There are 205 communes assigned to treatment. Of the 205, 67 were selected by the four currently participating Phase 2 competitors as communes they plan to work in (205 – 67 = 138, the number of observations in the 'non-selected treatment' group).
This exhibit displays characteristics related to key outcome variables, and characteristics for which the two groups had statistically significant differences. A global F test shows that the two groups are distinguishable with respect to all baseline characteristics (p=0.032).
*p<0.1 **p<0.05 ***p<0.01

Next, we interviewed competitors to understand and try to mimic their selection process. In coordination with the cooperative leaders in control communes, we selected 12, 20, 8, and 6 farmers who meet the criteria listed in Exhibit B-2 for each of the four competitors respectively, based on a 2018 list of farmers. We then randomly ordered each set and recruited farmers in order of the randomly sorted list until we achieved a sufficient number of responses.

The competitors selected farmers in the selected areas tended to be whomever happened to, pre-program, own sufficiently suitable land next to other similarly suitable farmers. We might expect competitors to select commercially oriented farmers with larger plots, but we did not find evidence of this sort of selection. Some farmers prior to the program had only grown food for home consumption. Farmers were also not selected based on large plot size. We do believe that cooperative leaders selected for plots that drain better, and so asked all farmers, including those in the comparison, to select the three plots that drain best for comparison to avoid plot-level selection bias.

Using the information from interviews with competitors and cooperative leaders, and also using the information about participating farmers' plot sizes, we stratified the recruitment of comparison farmers by competitor-selection-type. Exhibit B-2 illustrates the stratification we conducted *within each control cooperative* to select comparison farmers that 'match' the type of farmers that might be recruited if the cooperative had been assigned to the treatment group.

Exhibit B-2. Sample selection protocol for comparison farmers in control communes

Compe-titor	Percent of AgResults spring 2019 farmers	Number of farmers to sample in each control cooperative (total 23)	Area/neighbourhood selection criteria (All: neighbourhood is flat/high elevation, or otherwise easy for discharge of water)	Number of farmers, by rice plot size (all farmers required to have minimum 300 m^2 under rice)
i4	26%	6	Fields are near the road (all fields were near a road)	4 farmers with > 900 m^2 2 farmers with < 900 m^2
i5	42%	10	Farmer has worked with the company previously	5 farmers with > 900 m^2 5 farmers with < 900 m^2
i18	19%	4	Area has lots of rice farmers.	1 farmer with > 900 m^2 3 farmers with < 900 m^2
i23	14%	3	Can use a transplanter (most farmers did not own the transplanter)	0.5 farmer with > 900 m^2 2.5 farmers with < 900 m^2

We did not attempt to mimic farmer-level selection into the program because, based on our cooperative leader interviews, there was little farmer-level selection. Farmers usually did not refuse to participate, although some later dropped out (often an entire area might drop together). This may be because farmers in cooperatives often follow the cooperative leaders' schedule regardless – announcements for times to fertilise are broadcasted on speakers throughout villages – and the farmers found the promise of economic gain credible.

Annex C – Baseline balance

Exhibit C-1 and Exhibit C-2 display the average season and the spring balance between the treatment and control farmers in the RCT. It shows overall balance. The spring sample is less complete than the summer sample, owing to the challenges of the high rates of COVID-19 in a few areas in the spring of 2020. The summer balance tables are even more similar to the all-season balance table than the spring. Exhibit C-3 and Exhibit C-4 provide baseline balance for the matched comparison and also shows overall balance.

Exhibit C-1. RCT diary study: all-season balance test between farmers in treatment and control communes

Outcome	Treatment (A)	Control (B)	Differ-ence (A-B)	Standard Error	Signifi-cance (p-value)
Farmer characteristics					
Age	59.2	54.1	5.0	3.9	0.192
Male	80.2	72.4	7.7	5.7	0.174
Rice area owned (m2)*	4674.8	8606.8	-3932.0	4307.5	0.362
Percent of cooperative members who have farmers larger than 1500 square meters	47.7	45.2	2.5	3.1	0.421
Percent of population that are listed members of cooperative	37.9	43.0	5.1	4.4	0.243
Rice farming					
Average yield (MT/hectare)	7.0	7.0	0.0	0.1	0.806
Percent of commune area used for rice cultivation, Spring 2018	52.9	50.2	2.7	2.0	0.184
Average farmer rice cultivation area per cooperative (ha)	0.2	0.2	0.0	0.0	0.339
Rice field is at high elevation	15.4	11.4	4.0	3.3	0.228
Rice field is at low elevation	12.6	13.7	-1.1	3.5	0.758
Drain quality (1 best, 5 worst)	1.5	1.6	-0.2	0.1	0.059
Water managed solely by leader	29.3	39.3	-10.0	8.2	0.224
Cooperative has at least 1 riding transplanter	7.3	11.8	-4.5	5.3	0.395
Weather and geography					
Average precipitation, January through May (imputed)	69.5	69.3	0.2	0.3	0.594
Minimum temperature, January through May (imputed)	14.0	14.0	0.0	0.0	0.445
Soil organic carbon stock at depth 5 to 15cm (ton/ha), 250m resolution (imputed)	219.6	220.2	-0.5	7.9	0.946
Sum of evapotranspiration from January 2019 through May 2019 (mm) (imputed)	364.9	364.5	0.4	3.7	0.918
Average travel time from commune to city with 50K to 100K people (imputed)	9.8	9.2	0.6	1.1	0.559

Data source: Baseline commune-level administrative data and farmer diaries.
Notes: There are 177 communes in the treatment sample and 50 communes in the control sample. The estimates here are obtained using a sample from the diary study for which we have complete information on covariates and the outcome “amount of nitrogen applied per sao”. There are 359 and 380 farmers in the treatment sample in the spring and summer seasons, and 92 and 96 farmers in the control sample in the spring and summer.
*A few outliers in in the control group drive the large difference in means. The 50th, 75th, and 90th percentiles are more or less equal.

Exhibit C-2. RCT diary study: spring balance test between treatment and control communes

Outcome	Treatment (A)	Control (B)	Difference (A-B)	Standard Error	Significance (p-value)
Farmer characteristics					
Age	56.1	53.7	2.4	1.5	0.096
Male	82.2	72.0	10.2	5.8	0.078
Rice area owned (m2)	4775.4	8924.0	-4148.6	4477.9	0.355
Percent of cooperative members who have farmers larger than 1500 square meters	47.8	44.5	3.3	3.2	0.303
Percent of population that are listed members of cooperative	37.2	44.0	06.8	4.6	0.145
Rice farming					
Average yield (MT/hectare)	7.0	7.0	0.0	0.1	0.939
Percent of commune area used for rice cultivation, Spring 2018	52.8	49.8	2.9	2.1	0.158
Average farmer rice cultivation area per cooperative (ha)	0.2	0.2	0.0	0.0	0.239
Rice field is at high elevation	17.2	12.9	4.3	4.3	0.314
Rice field is at low elevation	10.6	15.6	-4.9	4.0	0.218
Drain quality (1 best, 5 worst)	1.5	1.7	-0.2	0.1	0.062
Water managed solely by leader	29.6	38.6	-9.0	8.4	0.284
Cooperative has at least 1 riding transplanter	8.5	13.4	-4.9	5.9	0.408
Weather and geography					
Average precipitation, January through May (imputed)	69.5	69.3	0.1	0.3	0.631
Minimum temperature, January through May (imputed)	14.0	14.0	0.0	0.0	0.551
Soil organic carbon stock at depth 5 to 15cm (ton/ha), 250m resolution (imputed)	219.5	221.6	-2.2	7.2	0.765
Sum of evapotranspiration from January 2019 through May 2019 (mm) (imputed)	365.3	363.5	1.8	3.8	0.642
Average travel time from commune to city with 50K to 100K people (imputed)	9.8	9.3	0.5	1.2	0.657

Exhibit C-3. Matched comparison study: all season balance test between treatment and comparison plots

Outcome	Treat-ment (A)	Con-trol (B)	Differ-ence (A-B)	Stan-dard Error	Signifi-cance (p-value)
Farmer characteristics					
Head of household is female	21.3	18.0	3.3	2.1	0.131
Proportion of farmers for whom rice is main source of income	37.7	35.3	2.4	3.1	0.445
Average rice plot size (m2)	2019.5	2042.6	-23.1	76.5	0.762
Number of types of cash crops grown	0.9	1.1	-0.2	0.1	0.165
Proportion of farmers who completed secondary school	90.3	87.6	2.7	1.7	0.108
Number of assets owned	3.0	2.9	0.1	0.1	0.247
Total types of animals owned	3.6	4.0	-0.4 *	0.2	0.066
Percent of population that are listed members of cooperative	42.6	39.8	2.8	5.6	0.616
Percent of cooperative members who have farmers larger than 1500 square meters	53.3	53.9	-0.6	3.0	0.835
Rice farming (cooperative level unless otherwise specified)					
Average value of production per farmer per cooperative (VND/MT)	9065.8	8935.1	130.7	473.3	0.783
Average yield (MT/hectare)	7.1	7.1	0.0	0.1	0.595
Cooperative has at least 1 riding transplanter	14.6	8.3	6.3	6.1	0.306
Percent of commune area used for rice cultivation, Spring 2018	52.9	54.0	-1.1	2.3	0.633
Average farmer rice cultivation area per cooperative (ha)	17.5	18.1	-0.6	0.9	0.511
At least one of top 3 common rice varieties at baseline is an approved AgResults	81.0	85.2	-4.2	7.6	0.585
Water managed solely by leader	23.0	28.8	-5.8	9.0	0.520
Irrigation system in place (yes=1; Not yet=0)	61.5	57.1	4.4	11.4	0.699
Weather and geography					
Sum of evapotranspiration from January 2019 through May 2019 (mm) (imputed)	365.4	377.2	-11.8 **	5.5	0.034
Minimum temperature, January through May (imputed)	14.1	14.1	0.0	0.0	0.827
Soil organic carbon stock at depth 5 to 15cm (ton/ha), 250m resolution (imputed)	211.4	220.5	-9.1	8.6	0.288
Average travel time from commune to city with 50K to 100K people (imputed)	8.9	8.6	0.3	1.1	0.816
Average precipitation, January through May (imputed)	68.9	68.6	0.3	0.5	0.469

Data source: Spring and summer income surveys.
Notes: There are 869 and 1044 farmers in the treatment sample in the spring and summer seasons, and 1197 and 1074 farmers in the control sample in the spring and summer. We included AgResults plots from the treatment farmers (1339 in spring, 1779 in summer) and all plots of the comparison farmers (2644 in spring, 2847 in summer). Clustering is at the village level. There are 73 and 97 villages in the treatment sample in the spring and summer seasons, and 88 and 85 villages in the comparison sample in the spring and summer season.

Exhibit C-4. Matched comparison study: spring balance test between treatment and comparison plots

Outcome	Treat-ment (A)	Con-trol (B)	Differ-ence (A-B)	Stan-dard Error	Signifi-cance (p-value)
Farmer characteristics					
Head of household is female	24.2	18.1	6.1 **	2.9	0.041
Proportion of farmers for whom rice is main source of income	47.0	43.5	3.5	3.6	0.330
Total rice paddies owned (m2)	2035.9	2146.4	-110.5	101.1	0.276
Number of types of cash crops grown	0.8	1.0	-0.2 *	0.1	0.074
Proportion of farmers who completed secondary school	90.7	87.6	3.1	1.9	0.111
Number of assets owned	2.9	2.9	0.0	0.1	0.992
Total types of animals owned	3.3	4.0	-0.7 ***	0.2	0.000
Percent of population that are listed members of cooperative	55.1	57.9	-2.8	3.6	0.446
Percent of cooperative members who have farmers larger than 1500 square meters	36.3	37.8	-1.5	5.7	0.795
Rice farming					
Average value of production per farmer per cooperative (VND/MT)	7.1	7.1	0.0	0.1	0.771
Average yield (MT/hectare)	9168.2	9547.3	-379.1	584.0	0.517
Cooperative has at least 1 riding transplanter	14.5	3.6	10.9	6.6	0.102
Percent of commune area used for rice cultivation, Spring 2018	53.7	54.4	-0.7	3.1	0.829
Average farmer rice cultivation area per cooperative (ha)	17.9	19.6	-1.7	1.2	0.163
At least one of top 3 common rice varieties at baseline is an approved AgResults	87.2	90.8	-3.6	6.7	0.598
Water managed solely by leader	20.4	19.4	1.0	8.5	0.902
Irrigation system in place (yes=1; Not yet=0)	60.7	59.3	1.4	14.4	0.923
Weather and geography					
Sum of evapotranspiration from January 2019 through May 2019 (mm) (imputed)	8.4	8.2	0.2	1.2	0.865
Minimum temperature, January through May (imputed)	364.5	382.1	-17.6 ***	6.0	0.004
Soil organic carbon stock at depth 5 to 15cm (ton/ha), 250m resolution (imputed)	69.1	68.1	1.0 *	0.6	0.077
Average travel time from commune to city with 50K to 100K people (imputed)	14.1	14.1	-0.0	0.0	0.543
Average precipitation, January through May (imputed)	212.4	217.4	-5.0	7.5	0.504

Data source: Spring income surveys.
Notes: There are 869 farmers in the treatment sample in the spring season, and 1197 farmers in the control sample in the spring. We included AgResults plots from the treatment farmers (1339 in spring) and all plots of the comparison farmers (2644 in spring). Clustering is at the village level. There are 73 villages in the treatment sample in the spring, and 88 villages in the comparison sample in the spring season.

Annex D – Extra exhibits pertaining to Evaluation Question 2

As expected, the AgResults competition did not have significant impacts on the uptake of any rice cultivation practices promoted by the AgResults competitors. Exhibit D-1 shows that responses of the farmers living in communes assigned to treatment (mostly not part of AgResults) were similar to the responses of farmers living in communes assigned to be control. The differences between the treatment communes and control communes have a small magnitude and none are statistically significant.

Exhibit D-1. RCT: Impact across Thai Binh of AgResults on technology package use

Outcome	Average across all of Thai Binh province	Communes assigned to treatment group (A)	Communes assigned to control group (B)	Difference (A-B)	Standard error	Significance (p-value)
Farmer fully complies with any one of the AgResults competitors' technology packages (%)						
Average, spring and summer	1.0	0.9	1.7	-0.8	0.9	0.341
Spring 2020	0.2	0.0	0.9	-0.9	0.9	0.328
Summer 2020	1.9	1.6	2.5	-0.9	1.7	0.590
Farmer uses AgResults-approved rice variety (%)						
Average, spring and summer	69.5	69.8	69.7	0.0	4.3	0.994
Spring 2020	54.6	53.8	55.6	-1.8	7.2	0.802
Summer 2020	84.4	85.1	83.9	1.2	5.5	0.829
Seed density is within range of AgResults recommendations (%)						
Average, spring and summer	37.3	37.0	36.1	0.9	5.7	0.878
Spring 2020	34.8	33.4	36.1	-2.6	7.1	0.715
Summer 2020	39.8	41.9	36.7	5.2	6.7	0.441
Number of fertiliser applications within range of AgResults recommendations (%)						
Average, spring and summer	70.1	71.8	68.3	3.5	5.5	0.526
Spring 2020	73.3	75.0	76.1	-1.1	6.0	0.850
Summer 2020	66.8	69.3	63.3[25]	6.0	6.8	0.376
Nitrogen application within range of AgResults recommendations (%)						
Average, spring and summer	70.0	69.9	71.2	-1.3	4.9	0.783
Spring 2020	67.9	68.7	66.8	1.9	6.6	0.773
Summer 2020	72.0	70.8	73.8	-3.0	5.6	0.597
Percent of days that field is dry is within range of AgResults recommendations (%)						
Average, spring and summer	30.6	30.5	32.4	-1.9	4.5	0.669
Spring 2020	27.7	25.4	34.1	-8.7	6.2	0.163
Summer 2020	33.6	37.8	32.5	5.3	6.8	0.439
Farmer used lime or bioenzyme to treat crop residue (%)						
Average, spring and summer	25.6	25.7	24.4	1.2	5.5	0.822
Spring 2020	27.1	27.1	27.6	-0.4	7.5	0.956
Summer 2020	24.0	24.9	22.4[26]	2.5	5.8	0.672

[25] There is less time to apply fertiliser in the summer than in the spring. Anecdotally, farmers believe if their crops receive nutrients from stubble they can apply less fertiliser.
[26] Farmers may have not applied yet as the next crop is still further out.

Outcome	Average across all of Thai Binh province	Communes assigned to treatment group (A)	Communes assigned to control group (B)	Difference (A-B)	Standard error	Significance (p-value)
Farmer burns crop residue (%)						
Average, spring and summer	35.8	36.7	34.7	1.9	4.7	0.682
Spring 2020	38.2	38.3	42.2	-3.9	7.1	0.582
Summer 2020	33.4	33.4	26.4[27]	6.9	6.0	0.251

Source: AgResults Independent Evaluator's Farmer Diary.
Note: There are 359 and 380 farmers in treatment communes in the spring and summer, respectively, and 92 and 96 farmers in control communes in spring and summer, respectively, with non-missing covariates and a non-missing outcome measure for nitrogen application, and it is on this sample that balance was assessed. They represent 175 and 177 treatment comments in the spring and summer, and 48 and 50 control communes in the spring and summer.
Means are regression-adjusted. The number of treatment and comparison plots in the spring is 2835 and 3669; summer 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

[27] At the time of data collection, the straw and stubble were still wet; farmers were waiting to burn until the crop was dry.

Exhibit D-2. Matched comparison study: Impact of AgResults participation in uptake of technology package components among AgResults farmers

Outcome	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	% Change (A-B)/B	Standard error	Significance (p-value)
Used AgResults-equivalent rice variety						
Average, spring and summer crop	90.0	52.8	37.2 ***	70.5	3.3	0.000
Spring 2020	82.7	44.9	37.8 ***	84.2	4.7	0.000
Summer 2020	95.6	61.2	34.4 ***	56.2	3.5	0.000
Planting density (kgs/sao)						
Average, spring and summer crop	1.4	1.5	-0.1 **	-6.7	0.0	0.011
Spring 2020	1.2	1.4	-0.2 ***	-14.3	0.0	0.000
Summer 2020	1.4	1.5	-0.1 ***	-6.7	0.1	0.008
Used fertiliser variety recommended by a competitor (%)						
Average, spring and summer crop	72.5	27.7	44.8 ***	161.7	2.8	0.000
Spring 2020	69.9	29.9	40.0 ***	133.8	4.0	0.000
Summer 2020	77.1	27.5	49.6 ***	180.4	3.5	0.000
Number of times apply fertiliser						
Average, spring and summer crop	3.0	2.7	0.3 ***	11.1	0.1	0.000
Spring 2020	2.9	2.8	0.1	3.6	0.1	0.366
Summer 2020	2.9	2.6	0.3 ***	11.5	0.1	0.000
Nitrogen applied (kgs/sao)						
Average, spring and summer crop	2.6	3.0	-0.4 **	-13.3	0.2	0.025
Spring 2020	2.7	3.2	-0.5 ***	-15.6	0.2	0.007
Summer 2020	2.7	2.8	-0.1	-3.6	0.1	0.321
Number of days the plot was completely dry						
Average, spring and summer crop	11.1	11.9	-0.8	-6.7	0.9	0.410
Spring 2020	22.8	17.7	5.1 ***	28.8	1.1	0.000
Summer 2020	3.6	5.5	-1.9 ***	34.6	0.6	0.001
Used bioenzymes on straw (%)						
Average, spring and summer crop	39.0	4.7	34.3 ***	729.8	3.5	0.000
Spring 2020	53.3	5.9	47.4 ***	803.4	4.8	0.000
Summer 2020	37.8	3.3	34.5 ***	1045.5	3.3	0.000

Outcome	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	% Change (A-B)/B	Standard error	Significance (p-value)
Burned straw (%)						
Average, spring and summer crop	35.20	41.10	-5.9 *	-14.4%	3.3	0.071
Spring 2020	41.90	49.60	-7.7 **	-15.5%	3.6	0.036
Summer 2020	32.00	33.20	-1.2	-3.6%	3.8	0.749

Source: AgResults Independent Evaluator's Farmer Income Survey.
Note: Means are regression-adjusted. The number of treatment and comparison farmer plots in the spring is 2835 and 3669; summer 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

Exhibit D-3. Matched comparison study: Gender-differentiated results of AgResults on technology component use

Outcome	Average, female-headed AgResults households (A)	Average, male-headed AgResults households (B)	Average, female-headed non-AgR households (C)	Average, Male-headed non-AgR households (D)	Female impact (A)-(C)	Male impact (B)-(D)	Differential impact [(A) – (C)] – [(B) – (D)]
Farmer uses AgResults-approved rice variety (%)							
Average, spring and summer	86.8	93.7	55.0	55.0	31.8	38.7	-6.9 **
Spring 2020	90.5	87.3	57.7	48.1	32.8	39.1	-6.3
Summer 2020	94.1	96.8	63.2	61.3	30.9	35.6	-4.7
Planting density (kg/sao)							
Average, spring and summer	1.4	1.3	1.5	1.5	-0.1	-0.1	0.0
Spring 2020	1.3	1.2	1.4	1.4	-0.2	-0.2	0.0
Summer 2020	1.3	1.3	1.4	1.5	-0.1	-0.2	0.0
Number of fertiliser applications							
Average, spring and summer	2.8	2.9	2.6	2.6	0.2	0.3	-0.1 **
Spring 2020	2.6	2.8	2.7	2.6	-0.1	0.1	-0.2 **
Summer 2020	2.8	2.9	2.5	2.6	0.2	0.4	-0.1 *
Nitrogen application (kg/sao)							
Average, spring and summer	2.6	2.7	2.7	3.1	-0.2	-0.4	0.2 **
Spring 2020	2.5	2.6	2.8	3.1	-0.4	-0.5	0.1
Summer 2020	2.7	2.8	2.6	2.9	0.1	-0.2	0.3 *
Number of days that field is dry							
Average, spring and summer	9.9	10.8	11.6	11.6	-1.6	-0.9	-0.8
Spring 2020	19.1	21.3	17.0	15.8	2.1	5.6	-3.4 **
Summer 2020	3.4	3.8	5.7	5.6	-2.3	-1.8	-0.5

Outcome	Average, female-headed AgResults households (A)	Average, male-headed AgResults households (B)	Average, female-headed non-AgR households (C)	Average, Male-headed non-AgR households (D)	Female impact (A)-(C)	Male impact (B)-(D)	Differential impact [(A) – (C)] – [(B) – (D)]
Farmer used bioenzymes to treat crop stubble (%)							
Average, spring and summer	69.0	65.3	11.9	11.7	57.2	53.6	3.6
Spring 2020	80.5	77.4	12.3	16.2	68.2	61.2	7.0 *
Summer 2020	74.4	68.8	11.6	6.3	62.8	62.5	0.3
Agree that most or all of the people who applied herbicide were women (%)							
Average, spring and summer	53.4	46.8	47.4	47.4	6.0	-0.6	6.6
Spring 2020	84.9	48.8	74.2	45.2	10.7	3.6	7.1
Summer 2020	70.1	38.0	62.3	38.1	7.8	-0.2	8.0
Agree that most or all of the people who applied pesticide were women (%)							
Average, spring and summer	56.9	32.4	60.9	30.7	-4.0	1.7	-5.7
Spring 2020	74.7	40.0	73.0	38.0	1.7	2.0	-0.3
Summer 2020	47.1	29.3	50.2	25.0	-3.1	4.3	-7.4
Agree that most or all of the people who did the land prep and planting were women (%)							
Average, spring and summer	80.1	45.9	86.8	38.4	-6.6	7.5	-14.1 ***
Spring 2020	75.7	36.6	82.6	32.2	-6.9	4.4	-11.2 *
Summer 2020	79.1	46.4	86.5	41.4	-7.5	5.0	-12.4 **
Agree that most or all of the people who pulled rice seedlings were women (%)							
Average, spring and summer	97.7	83.1	97.4	76.7	0.3	6.4	-6.2 *
Spring 2020	97.9	83.0	97.6	79.2	0.3	3.8	-3.5
Summer 2020	97.7	81.6	96.0	74.6	1.7	7.0	-5.3
Agree that most or all of the people who seeded or transplanted were women (%)							
Average, spring and summer	89.8	68.0	92.1	62.3	-2.3	5.7	-8.0 **
Spring 2020	93.4	67.4	95.7	67.3	-2.3	0.1	-2.5
Summer 2020	88.3	68.1	89.1	57.2	-0.8	11.0	-11.8 **
Agree that most or all of the people who did irrigation and draining were women (%)							
Average, spring and summer	81.6	47.8	83.5	33.5	-1.9	14.3	-16.2 ***
Spring 2020	70.6	50.2	97.9	29.5	-27.3	20.7	-48.0 ***
Summer 2020	86.0	45.0	73.6	34.6	12.4	10.3	2.0

Outcome	Average, female-headed AgResults households (A)	Average, male-headed AgResults households (B)	Average, female-headed non-AgR households (C)	Average, Male-headed non-AgR households (D)	Female impact (A)-(C)	Male impact (B)-(D)	Differential impact [(A) – (C)] – [(B) – (D)]
Agree that most or all of the people who did the weeding were women (%)							
Average, spring and summer	94.8	77.3	97.9	75.8	-3.1	1.5	-4.7
Spring 2020	96.5	78.8	95.8	75.6	0.7	3.2	-2.5
Summer 2020	92.0	74.8	99.3	75.6	-7.3	-0.8	-6.5

Source: AgResults Independent Evaluator's Income Survey.

Note: Means are regression-adjusted. In the spring season, there were 668 and 980 male-headed households in the AgResults and comparison group, respectively; and 211 and 217 female-headed households in the AgResults and comparison group. In the summer season, there were 844 and 876 male-headed households in the AgResults and comparison group, respectively; and 222 and 198 female-headed households in the AgResults and comparison group.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

Annex E – Extra exhibits pertaining to Evaluation Question 3

Exhibit E-1. Impact of technology package uptake on revenue

Outcome	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	Percent change =(A-B)/A	Standard error	Significance (p-value)
Net value – value if sold all harvested rice at sales price minus costs ('000 VND/sao)						
Average, spring and summer crop	887.3	802.3	85.0 **	10.6	35.2	0.017
Spring 2020	1062.7	859.6	203.1 ***	23.6	45.2	0.000
Summer 2020	842.7	761.4	81.3 *	10.7	41.6	0.052
Gross sales revenue ('000 VND/sao)[a]						
Average, spring and summer crop	898.8	550.5	348.3 ***	63.3	58.2	0.000
Spring 2020	1241.4	744.9	496.5 ***	66.7	80.8	0.000
Summer 2020	794.7	360.1	434.6 ***	120.7	54.2	0.000
Net sales revenue (income minus costs) ('000 VND/sao)						
Average, spring and summer crop[a]	270.0	-75.9	345.9 ***	455.7	57.7	0.000
Spring 2020	588.1	110.2	477.9 ***	433.7	78.8	0.000
Summer 2020	180.0	-252.5	432.5 ***	171.3	54.9	0.000
Total costs ('000 VND/sao)						
Average, spring and summer crop	568.0	580.3	-12.3	-2.1	12.5	0.326
Spring 2020	579.0	584.9	-5.9	-1.0	17.9	0.740
Summer 2020	561.8	571.5	-9.7	-1.7	14.0	0.490
Rice sale price ('000 VND)						
Average, spring and summer crop	7.5	8.0	-0.5 ***	-6.3	0.1	0.000
Spring 2020	7.8	8.1	-0.3 *	-3.7	0.2	0.070
Summer 2020	7.6	8.0	-0.4 **	-5	0.2	0.023

Outcome	AgResults plots of AgResults farmers (A)	All plots of comparison farmers (B)	Difference (A-B)	Percent change =(A-B)/A	Standard error	Significance (p-value)
Proportion of farmers that sell rice (yes/no)						
Average, spring and summer crop	50.0	38.0	12.0 ***	31.6	2.8	0.000
Spring 2020	53.6	46.2	7.4 **	16.0	3.6	0.041
Summer 2020	50.1	29.6	20.5 ***	69.3	2.9	0.000
Amount of rice sold (kgs)						
Average, spring and summer crop	388.1	281.9	106.2 ***	37.7	32.3	0.001
Spring 2020	484.6	372.3	112.3 ***	31.2	36.2	0.002
Summer 2020	361.4	182.4	179.0 ***	98.1	33.5	0.000

Source: AgResults Independent Evaluator's Farmer Income Survey.
Note: Means are regression-adjusted. The number of treatment and comparison plots in the spring is 2835 and 3669; summer 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.
[a] For sales revenue minus costs, the average difference for both seasons combined (345) is lower than either season individually (484 for spring, 428 for summer). Close examination reveals this as a logical consequence of averaging the two seasons' sales revenue for the AgResults farmers (488 and 174 average to 227, roughly due to regression adjustment) and averaging the seasons' sales revenue for the comparison farmers (5 and -231 average to -118, roughly due to regression adjustment).

Exhibit E-2. Imputed impact of technology package uptake on revenue in the absence of AgResults-associated competitor incentives

Outcome	Mean of farmers' AgResults plots (A)	Imputed mean in absence of competitor incentives (B)	Mean of all comparison farmer plots (C)	Difference (B-C)	% Change (B-C)/C	Standard error
Net value ('000 VND/sao)						
Average, spring and summer	887.3	828.6	802.3	26.3	3.3	26.9
Spring 2020	1062.7	935.7	859.6	76.1 *	8.9	39.7
Summer 2020	842.7	790.9	761.4	29.5	3.9	33.4
Net sales revenue (income minus costs) ('000 VND/sao)						
Average, spring and summer crop	270.0	198.6	-75.9	274.5 ***	361.7	51.9
Spring 2020	588.1	472.1	110.2	361.9 ***	328.4	76.7
Summer 2020	180.0	110.6	-252.5	363.1 ***	143.8	47.3
Total costs ('000 VND/sao)						
Average, spring and summer	568.0	608.1	580.3	27.8 **	4.8	12.2
Spring 2020	579.0	616.6	584.9	31.7 *	5.4	16.6
Summer 2020	561.8	606.4	571.5	34.9 **	6.1	14.0

Source: AgResults Independent Evaluator's Farmer Income Survey.
Note: Means are regression-adjusted. The number of treatment and comparison plots in the spring is 2835 and 3669; summer 3186 and 3270.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.

Exhibit E-3. Average costs and revenue outcomes in Spring 2020, by competitor and technology package

Competitor	Variety	Number of observations	Seed cost per sao ('000 VND/sao)	Fertiliser cost per sao ('000 VND)	Total cost per sao ('000 VND/sao)	Expected cost absent AgResults ('000)	Yield (MT/ha)	Proportion who sold rice	Rice sale price per kg ('000 VND)	Rice value per sao ('000 VND/sao)[a]
i4	DS1	180	38	191	628	647	6.43	96%	6.31	780
i4	DS3	49	42	167	652	682	5.19	80%	6.41	479
i4	BC15	125	37	160	691	717	6.00	75%	7.42	902
i5	BC15	416	33	151	651	715	6.46	89%	7.99	1322
i5	BT7	0	.	.	.	.	.	.	.	.
i18	LTH31	144	21	163	663	727	5.65	45%	7.30	810
i18	BT7	187	32	142	565	598	4.80	46%	7.72	753
i18 (not a spring technology package)	BC15	8	36	223	652	738	5.02	75%	6.73	648
i23	T10	85	28	181	633	746	4.73	39%	8.60	860
i23	BT7	90	32	181	601	679	4.18	28%	8.36	666
i23	DT8	26	37	152	571	750	4.68	62%	8.55	869
i23 (not a spring technology package)	BC15	15	30	166	691	773	5.35	47%	7.09	822
Comparison	DS1	3	36	224	665	665	6.61	67%	5.93	720
Comparison	DS3	0	.	.	.	.	.	.	.	.
Comparison	BC15	281	42	166	635	635	5.19	50%	7.36	781
Comparison	BT7	569	35	178	610	610	4.40	38%	8.25	723
Comparison	LTH31	0	.	.	.	.	.	.	.	.
Comparison	T10	330	36	171	606	606	4.21	35%	9.11	761
Comparison	DT8	125	43	167	606	606	4.50	38%	7.84	740

Note: Please interpret these numbers with caution. The confidence intervals here are wide. Also, our study did not have the power to detect differences in many comparisons, and so comparisons between numbers here may not be accurate.

Exhibit E-4. Average costs and revenue outcomes in Summer 2020, by competitor and technology package

Competitor	Variety	Number of observations	Seed cost per sao ('000 VND/sao)	Fertiliser cost per sao ('000 VND/sao)	Total cost per sao ('000 VND/sao)	Expected cost absent AgResults ('000 VND)	Yield (MT/ha)	Proportion who sold rice	Rice sale price per kg ('000 VND)	Rice value per sao ('000 VND/sao)
i4	DS1	293	36	212	656	697	5.62	97%	8.07	680
i4	DS3	3	39	196	631	639	5.56	100%	10.17	669
i4	BC15	98	37	172	650	693	5.27	39%	8.22	644
i5	BC15	408	35	145	619	688	5.27	80%	8.11	1148
i5	BT7	20	36	117	533	635	4.48	90%	10.09	1030
i18	LTH31	0	.	.	.	.	.	.	.	.
i18 (not a summer technology package)	BT7	0	.	.	.	.	.	.	.	.
i18	BC15	681	36	139	589	630	5.38	44%	7.01	781
i23 (not a summer technology package)	T10	23	32	157	574	689	4.30	4%	10.16	931
i23	BT7	38	29	169	589	690	4.43	21%	10.12	1050
i23 (not a summer technology package)	DT8	17	46	170	560	653	4.69	41%	8.55	892
i23	BC15	218	40	179	636	710	4.82	34%	7.88	799
Comparison	DS1	40	43	190	627	627	5.43	88%	6.53	651
Comparison	DS3	1	60	143	305	305	6.53	100%	8.81	1316
Comparison	BC15	1450	46	160	619	619	4.62	27%	7.84	713
Comparison	BT7	154	35	191	573	573	4.29	21%	9.14	949
Comparison	LTH31	0	.	.	.	.	.	.	.	.
Comparison	T10	71	37	198	631	631	4.43	25%	9.41	879
Comparison	DT8	89	41	165	625	625	4.59	20%	8.18	721

Note: This is by the number of plots, not number of farmers. The number of treatment and comparison plots in the spring is 2835 and 3669; summer 3186 and 3270. Some plots are missing information for some of the variables reported. Please interpret these numbers with caution. The confidence intervals here are wide. Also, our study did not have the power to detect differences in many comparisons, and so comparisons between numbers here may not be accurate.

Exhibit E-5. Gender-differentiated impacts of AgResults on income and costs, comparing AgResults farmers to matched comparison farmers

Outcome	Average, female-headed AgResults households (A)	Average, male-headed AgResults households (B)	Average, female-headed non-AgR households (C)	Average, male-headed non-AgR households (D)	Female impact (A)-(C)	Male impact (B)-(D)	Differential impact [(A) – (C)] – [(B) – (D)]
Net value ('000 VND/sao)							
Average, spring and summer	944.1	868.2	798.8	798.8	145.3	69.4	75.8 **
Spring 2020	1035.6	1036.9	823.0	829.8	212.7	207.1	5.6
Summer 2020	944.9	827.4	756.5	769.4	188.4	58.1	130.3 ***
Gross sales revenue ('000 VND/sao)							
Average, spring and summer	909.2	874.1	528.2	528.2	381.0	345.9	35.1
Spring 2020	1176.1	1169.2	649.1	654.9	526.9	514.2	12.7
Summer 2020	827.8	809.3	376.9	375.9	451.0	433.3	17.6
Net sales revenue (income minus costs) ('000 VND/sao)							
Average, spring and summer	255.3	254.4	-96.3	-96.3	351.6	350.8	0.9
Spring 2020	502.8	534.7	41.8	23.0	461.0	511.8	-50.8
Summer 2020	208.4	193.3	-229.5	-240.0	437.9	433.3	4.6
Total costs ('000 VND/sao)							
Average, spring and summer	566.4	564.0	553.9	583.5	12.5	-19.4	32.0 **
Spring 2020	584.2	561.4	549.8	584.0	34.4	-22.5	57.0 ***
Summer 2020	553.5	564.9	554.3	576.0	-0.8	-11.1	10.3
Rice sale price ('000 VND)							
Average, spring and summer	7.7	7.5	8.0	8.0	-0.3	-0.5	0.3 **
Spring 2020	7.8	7.8	8.1	8.1	-0.3	-0.3	0.0
Summer 2020	8.0	7.6	8.0	8.0	0.0	-0.5	0.4 **
Proportion of farmers that sell rice (yes/no)							
Average, spring and summer	47.7	49.5	36.8	37.1	10.9	12.4	-1.5
Spring 2020	49.3	51.4	43.9	43.2	5.4	8.2	-2.8
Summer 2020	48.8	51.2	29.8	30.2	19.1	21.0	-1.9

Outcome	Average, female-headed AgResults households (A)	Average, male-headed AgResults households (B)	Average, female-headed non-AgR households (C)	Average, male-headed non-AgR households (D)	Female impact (A)-(C)	Male impact (B)-(D)	Differential impact [(A) – (C)] – [(B) – (D)]
Amount of rice sold (kgs)							
Average, spring and summer	355.0	384.8	248.6	278.6	106.4	106.2	0.2
Spring 2020	399.2	448.7	333.9	323.0	65.3	125.7	-60.4
Summer 2020	359.1	376.3	156.2	201.4	202.9	174.9	28.0
Yield (MT/ha)							
Average, spring and summer	5.5	5.5	4.8	4.8	0.8	0.6	0.1 *
Spring 2020	6.0	5.8	4.9	4.9	1.1	0.9	0.2 *
Summer 2020	5.3	5.3	4.6	4.7	0.7	0.6	0.1

Source: AgResults Independent Evaluator's Income Survey.
Note: Means are regression-adjusted. In the spring season, there were 2004 and 2940 plots for which we have data for male-headed households in the AgResults and comparison group, respectively; and 633 and 651 plots for which we have data for female-headed households in the AgResults and comparison group. In the summer season, there were 2532 and 2628 plots for which we have data for male-headed households in the AgResults and comparison group, respectively; and 666 and 594 plots for which we have data for female-headed households in the AgResults and comparison group.
***/**/* implies significance at the .01, .05, and .10 levels, respectively.