Bangladesh

The Global Agricultural and Food Security Program (GAFSP)

Integrated Agricultural Productivity Project (IAPP)

Impact Evaluation Concept Note

5 October, 2012

Contents

List of acronyms

BADC – Bangladesh Agricultural Development Corporation

DLS – Department of Livestock Services

DAE – Department of Agricultural Extension

DANIDA – Danish International Development Agency

DIME – Development Impact Evaluation Initiative

DoF – Department of Fisheries

FFS – Farmer Field School

GAFSP – The Global Agricultural and Food Security Project

IAPP – Integrated Agricultural Productivity Project

ICC – Intracluster Correlation

IFPRI – International Food Policy Research Institute

IPA – Innovations for Poverty Action

IPM – integrated pest management

IE – Impact Evaluation

MDES – minimum detectable effect size

MDG – Millennium Development Goal

PIU – Project Implementation Unit

USAID – United States Agency for International Development

1. Introduction

Over the last two decades, Bangladesh has achieved impressive growth and poverty reduction. Its agricultural sector grew at a rate of 4.8 percent between 1990 and 2005. But poverty-related food insecurity is widespread, bolstered by the soaring prices of key staples. The country has a poverty rate of over 30% and the highest incidence of malnutrition of all countries: in 2008, Bangladesh's food insecure population was estimated at 65.3 million.[1] The Government of Bangladesh is pushing for increased use of technology and more intensive agricultural practices to improve food security and sustain economic growth. In 2009, the Bangladeshi Government expanded its social safety net programs and allocated US$500 million in stimulus packages to support its agriculture among other sectors.

The Global Agriculture and Food Security Project (GAFSP) sponsors the Integrated Agricultural Productivity Project (IAPP) in Bangladesh, which is designed to develop new technologies and boost adoption through the farmer field schools approach (FFS).

The Impact Evaluation (IE) of the IAPP project will contribute to understanding the drivers of technology adoption through two lenses. First, the overall project approach will be evaluated using uses a randomized phase-in of project villages. This will be referred to as the "Overall Project Evaluation." Second, innovations will be tested to understand, within the approach, what mechanisms can deliver higher results. We will refer to this as the "Demonstration Plot Evaluation"

The Demonstration Plot Evaluation is designed to test a fundamental question about technology adoption: to what extent can "learning by doing" increase technology adoption over "learning by observing"? To answer this question, we will explore methods to improve technology adoption in farmer groups by comparing the relative effectiveness of single demonstration plots (the standard approach) to more distributed demonstration strategies which allow more people to experiment with new technology. This will be a randomized controlled trial, assigning different approaches to demonstration to different farmer groups. This IE will help the government understand how to best organize demonstration within a FFS, providing rigorous evidence on what approach leads to the highest level of technology adoption. It is designed to provide actionable results early in the project, allowing the government to incorporate its findings into the program implementation.

This impact evaluation is led by the World Bank's Development Impact Evaluation Initiative (DIME), the South Asia Agricultural Development team (SASDA), and the Government of Bangladesh's IAPP project implementation unit. It is in collaboration with and external research partners, the Yale University School of Management and the NGO Innovations for Poverty Action.

[1] Food and Agricultural Organization of the United Nations (FAO) and World Food Program (WFP). 2008. "FAO/WFP Crop and Food Supply Assessment Mission to Bangladesh".

2. The IAPP Project

The IAPP project is designed to improve the income and livelihoods of crop, fish, and livestock farmers in Bangladesh. It consists of four separate components:

1. Component 1: Technology Generation and Adaptation
2. Component 2: Technology Adoption
3. Component 3: Water Management
4. Component 4: Project Management

After consultations with the government and the Bank teams, it was decided that the impact evaluation would concentrate on the Components 2 and 3, which promote the adoption of more productive agricultural technology (including irrigation).[2]

The technology adoption component is comprised of three sub-components which will be promoted:

1. **Crops:** The Department of Agricultural Extension (DAE) will promote the use of new seeds and farming practices. These include improved rice varieties, vegetable production, legume production, farmyard manure, and green manure.
2. **Fisheries:** The Department of Fisheries (DoF) will promote new breeds and more intensive fish cultivation. Four breeds will be promoted: mono-sex tilapia, rui, thai koi, and pangas. Semi-intensive cultivation, including fertilization and feeding will be introduced.
3. **Livestock:** The Department of Livestock Services (DLS) will promote improved livestock management practices. These include goat vaccination, backyard poultry production, and improved dairy milk production.

IAPP will promote technology adoption through the Farmer Field School approach (FFS). FFS involves forming groups of farmers who meet bi-weekly to discuss their most important challenges in farming and work with extension agents to develop solutions to these problems. These groups become an important venue to promote and diffuse new technology. Farmer groups will consist of "Demonstration Farmers", who receive subsidies from IAPP to demonstrate new technologies, and "Adoption Farmers", who consist of the rest of the group and are encouraged to adopt the new technology with limited subsidies (but will receive training and guidance in the technology as part of the FFS).

FFS represents represent a major shift from the traditional agricultural extension approach practiced in Bangladesh, commonly referred to as the training-and-visit system (T&V).[3] Under T&V,

[2] Since the IE will be conducted early in the lifetime of the project, Components 2 and 3 will be promoting previously-developed technologies. Technologies developed by Component 1 should be part of the adoption efforts of IAPP, but only in later years.

extension agents meet with small groups of farmers to introduce new technologies, and urge these farmers to spread their experiences to their neighbors. FFS hopes to provide farmers with not just new technologies, but the knowledge to become informed and continually improving producers.[4] Under FFS, farmers are encouraged to interact with extension agents to develop customized solutions to their problems, and work to best adapt new technologies to their specific conditions. In theory, FFS will be more effective at technology adoption since it teaches farmers to be critical users of technology, as opposed to just adopting based on the urging of extension officials.

The project will take place in eight districts: Rangpur, Kurigram, Nilfamari and Lalmonirhat districts in the North and Barisal, Patuakhali, Barguna and Jhalokathi districts in the South. The project has selected 375 unions (sub-districts) which will receive project activities. IAPP expects to reach around 300,000 beneficiaries.

The Overall Project Evaluation will measure the effects of Components 2 and 3 of IAPP, with special focus on the crops and fisheries sub-components. The Demonstration Plot Evaluation will focus on the crops sub-component.

3. The Demonstration Plot Evaluation

3.1 Motivation

The IAPP promotes a very ambitious change in the way extension services are implemented in Bangladesh. The DIME-GAFSP collaboration offers a unique opportunity to use impact evaluation (IE) to take this learning experience to the next level and rigorously assess the benefits of different models. The proposed IE is closely aligned with the project objectives, and will allow the team to find out which approach to demonstration leads to higher rates of technology adoption within the farmers field school model.

(1) Standard demonstration plots work by transferring certain types of knowledge about a new production process to farmers. Primarily, this is information about the availability of the demonstrated crop and an example of yields *under certain conditions* on the plot of the demonstration farmer. However, farmers who are thinking about adopting a new farming process don't know how these yields that they are observing compare to yields they would themselves receive. These differences could be due to differences in soil quality, input usage, cultivation knowledge, etc. In fact, it is well documented that yields on farmer's fields in Bangladesh rarely approach the yields on

---

[3] Picciotto, R (1997). "Reconsidering agricultural extension". *The World Bank research observer (0257-3032)*, 12 (2), p. 249.

[4] Gotland, Erin M., Elisabeth Sadoulet, Alain De Janvry, Rinku Murgai, and Oscar Ortiz. 2004. "The Impact of Farmer Field Schools on Knowledge and Productivity: A Study of Potato Farmers in the Peruvian Andes." *Economic Development and Cultural Change* 53, no. 1:63–92.

demonstration plots.[5] If demonstration plots do not provide a realistic indication of potential yields from a new technology, this is likely to affect their ability to promote technology adoption. Additionally, it might result in a situation where farmers adopt crops ill-suited to their land, resulting in welfare loss.

One way to overcome this problem may be to simply have **(2) more demonstration farmers**: if farmer group members see more of their neighbors becoming successful growing a new crop[6], they are more likely gain accurate information about their chance of success. Furthermore, this allows more member of the farmer group to 'learn by doing', possibly making them more likely to continue growing the new crop. In a study on technology adoption during the green revolution in India, Foster and Rosenszweig[7] find that farmers' own experiences and that of their neighbors are important drivers of technology adoption and income.

Perhaps the greatest effect of 'demonstration' could even come from **(3) complete decentralization**. Under this model all members of the farmer group are encouraged to cultivate small 'demonstration' plots on their own land, essentially moving from 'learning by observing' to 'learning by doing'. In this case, all participating farmers would have an opportunity to learn how to cultivate the new crop, and would get a more accurate measure of what the yields on their own farm would be. But demonstration plots are costly to support, requiring the project to invest in seeds fertilizer, advice, and other inputs. Given a fixed amount of funding, increasing the amount of demonstration farmers requires having smaller plots, and potentially giving up on economies of scale. It's not clear what the optimal number of demonstration farmers is. In addition, farmers may need some additional incentives to participate in this scheme, given that they are not yet confident that the new crop will be an improvement over their old practices.

With these concerns in mind, the IAPP and DIME are planning to evaluate the relative effectiveness of three different demonstration approaches: standard demonstration plots, shared demonstration plots, and self-demonstration. Results from this evaluation will be rapid, with a survey in 2014 that will measure which approach led to the greatest adoption of the new seeds. These results can be immediately fed into the IAPP strategy for new and existing farmer groups, and can also be used to improve the design of future projects.

3.2 Evaluation Question and Description of Demonstration Approaches

The Demonstration Plot Evaluation attempts to test which approach to crop demonstration will cause the most farmers to adopt improved technologies in the following season. The three different demonstration approaches tested are:

[5] Sattar, Shiekh A. "Bridging the Rice Yield Gap in Bangladesh". In Bridging the Rice Yield Gap in the Asia-Pacific Region. By Minas K. Papdemetriou, Frank J. Dent and Edward M. Herath. Food and Agricultural Organization of the Uniiter Nations Regional Office for Asia and the Pacific. Bangkok, Thailand. October 2000.

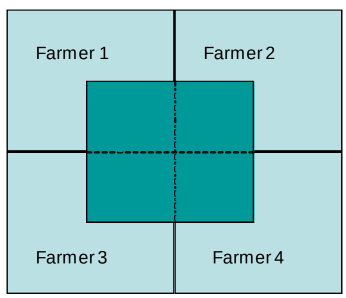
[6] Note that this "new crop" can be thought of as a different crop or simply a new variety of a previously cultivated crop.

[7] Rosenzweig, Mark R. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture". University of Chicago Press. Journal of Political Economy, Vol. 103, No. 6 (Dec., 1995), pp. 1176-1209

1. **Regular Demonstration plots**- This is the status quo in IAPP. For each type of technology introduced into the group (1-4 crops), one demonstration farmer is chosen. These demonstration farmers receive a 'package' of free seeds, fertilizer, and training. The selected farmers cultivate the promoted crop in the first year, and the rest of the group is expected to learn from this experience. In the second year, the rest of the farmers are encouraged to grow the crop. Farmers that adopt the technology in the second year receive free seeds but no inputs or special training from the project.

2. **Shared Demonstration Plots**- In this intervention, each demonstration 'package' (seeds, fertilizer and trainings) will be shared by two to four group members, as opposed to just one in the standard IAPP project. Where possible, the selected farmers will create demonstration plots on contiguous patches of land (see Figure 1 for a schematic), and they will be encouraged to work together to capture economies of scale. As in the demonstration plot intervention, demonstration farmers will receive free seeds, free inputs and trainings, but these resources will be spread over more farmers.

Figure 1: Shared Demonstration Plot. Dark green represents improved seed production.



3. **Incentives for Self-demonstration**- In this intervention, all members of the farmer field group are given the opportunity to grown the promoted variety in the first year, and the inputs that would be used on demonstration plots are instead spread out over all farmers who wish to participate. Farmers will be encouraged to grow the new crop on a small patch of land to test it out. Farmers who agree to grow the new crop in the first year will also receive an additional incentive: if the promoted variety does not perform as well as the old variety, they will receive a small cash payment (1000 taka, around $12.3 USD). The primary purpose of this payment is to send a signal to the farmers that the extension providers are confident that the new seed will perform better than the old. In order to see whether the payment should be applied, at the beginning of the season each participating farmer will pick a neighbor growing an older variety of the crop to be a reference farmer. If output on the reference farm is lower than output of the promoted variety, the farmer would receive his

small payment.[8] These payments would be made by DIME's research partner, the NGO Innovations for Poverty Action (IPA) using their own core research funding for Bangladesh.

3.3 Demo Plot Evaluation Design

The technology adoption experiment will be evaluated using a randomized controlled trial in two districts, Rangpur and Barisal. Within these districts, 220 villages will take part in the evaluation.

Before the cropping season, these villages will be randomly allocated into five treatment arms:

1. **Long Term Control (20 villages)**: Standard project activities (demonstration plots) beginning the final year of the project. Until then, they will have no project activities, and will just receive normal extension services from the government (roughly corresponding to the T&V system described above.)
2. **Short Term Control (36 villages)**: These villages will have standard project activities (demonstration plots) beginning in 2014. Until then, they will have no project activities, and will just receive normal extension services from the government.
3. **Demonstration Plots (54 villages):** These villages will have the standard IAPP project activities beginning in 2012.
4. **Shared Demonstration Plots (56 villages):** These villages will have demonstration plots shared among multiple farmers, as described above. These villages will start project activities in 2012.
5. **Incentives for self-demonstration (54 villages):** Instead of demonstration plots, all farmer group members will be offered an incentive to adopt the new crop variety, as described above. These villages will start project activities in 2012.

The short-term impact of the various treatment arms on variables of interest will be captured by comparing the outcome variables of each treatment group with both control groups, with data taken right before the project is rolled out in the short term control villages in 2014. Long-term impact will be determined with another round of data collection before the project is rolled out in the long term control villages in 2016.

3.4 Sampling

This section contains the sampling strategy for the Demonstration Plot Evaluation. This evaluation is taking place in two project districts, Rangpur and Barisal, as a pilot to inform scale up. Rangpur and Barisal were selected due to their high crop activity, and to give representation of one village from the North and one from the South.

[8] Note that this measurement will be done during the seeding phase of the plant, which gives a good prediction of the harvest, and will be conducted by IPA under supervisions of DIME. For data analysis purposes, yields will be measured post-harvest using household surveys. Since the surveys are not tied to the payouts, there should be no incentive to mis-report. Additionally, farmers will have to sign a contract saying they will cultivate the new crop to the best of their abilities, and this will be monitored by the FFS. To the extent that it is observable, farmers will not be able to receive a payout if they purposefully try to obtain poor yields on their demonstration plots.

3.4.1 Village Sampling

For the evaluation, we sampled 220 villages, 110 from Rangpur and 110 from Barisal. The village sampling strategy was as follows.

1. Start from the list of all villages eligible to receive the crop component in 2012.
2. Select a random sample of 220 villages, 110 from each district.
3. Randomize the 220 groups into the five treatment arms described above, stratifying by district.

This randomization was done by the DIME team using the randomized number generator in STATA, and was approved by the PIU.

3.4.2 Individual Sampling

For the Demo Plot Evaluation individuals were sampled for the baseline survey after farmer groups were formed. In each farmer group, 15 members were randomly chosen to be surveyed from the complete list of all members of the farmer group.

3.4.3 Power Calculations

In order to inform the sampling strategy, power calculations were undertaken using data from a recent USAID agricultural survey conducted by the International Food Policy Research Institute (IFPRI). This survey data was ready for one district that is part of the IAPP project (Barisal) which was used to conduct the power calculations. For calculating the predictive power at the baseline a panel survey is required, so we used data from an agricultural panel in India.[9]

The main outcome variable for the Demonstration Plot Evaluation will be the number of farmers who adopt and sustain use of the promoted crops. As the demonstration techniques are new, we don't have an accurate prediction of the effect size. (Note that we will be using a linear probability model to detect effects.)

Given these constraints, we used the data on paddy yield from our reference data to calculate intra-cluster correlation, and then settled on a sample size that resulted in a reasonably low minimum detectable effect size (MDES) given economic and logistical constraints.[10] As shown in Table 1, the MDES between any two treatment arms is .3

[9] The R^2 value calculated using total agricultural revenue from farmers in Gujarat, India. Survey described in: Cole, Shawn Allen, Giné, Xavier, Tobacman, Jeremy Bruce, Townsend, Robert M., Topalova, Petia B. and Vickery, James I., Barriers to Household Risk Management: Evidence from India (April 11, 2012). Harvard Business School Finance Working Paper No. 09-116

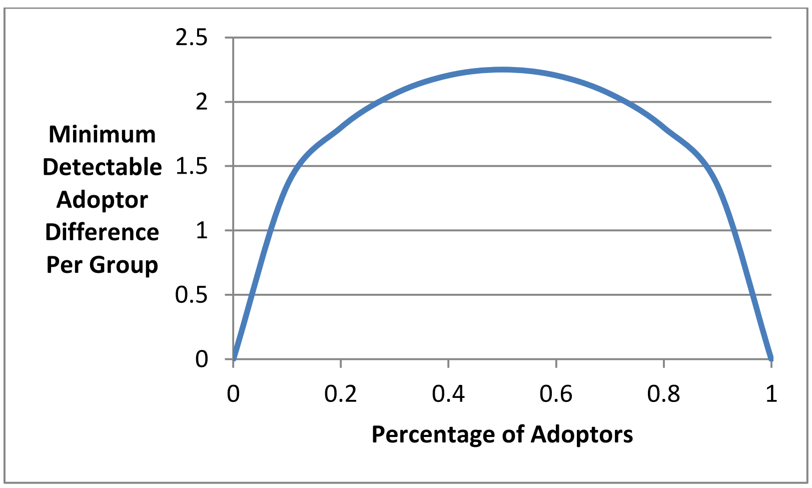
[10] The minimum detectable effects size is the minimum difference between two treatment arms that can be detected with a given power, normalized in standard deviations of the outcome variable. Therefore, an MDES of .3

Table 1: Power Calculation Data

Number of Villages	55
People Sampled Per Village	15
Intra-Cluster Correlation	.146
R^2 from baseline	.5
Size (α)	.05
Power	.85
MDES	.3

How can we interpret this MDES of .3? Since the outcome variable is binary, we can calculate its standard deviation for all values, and then calculate the difference in adoption to detect an effect of .3 for all possible average values of the outcome variable. This is presented in Figure 2 below:

Figure 2: Interpretation of MDES of .3



As Figure 2 shows, this MDES of .3 will allow us to detect relatively small effects in differences between the number of adopters in each treatment arm. Even in the situation where the effect is hardest to discern (which is where half the sample adopts), we will need an average difference of just over two adopters (per group of 15) to detect effects between treatment arms.

4. The Overall Project Evaluation

means that we can detect a difference between any two treatment arms of .3 standard deviations of the mean of the outcome variable.

The Overall Project Evaluation measures the overall effect of IAPP activities on farmer livelihood. Its primary focus is on crops and fisheries groups, while also attempting to measure the effect of livestock and water management activities.

4.1 Motivation

Although Bangladesh has seen increases in agricultural productivity over the last decades, its farmers are still producing far below potential. The estimated yield gap for paddy corresponds to a potential yield increase of 24% and 55% for the boro and aus seasons respectively.[11][12] Additionally, there is much opportunity to increase fish yields; in 2005/06 Bangladeshi fish farmers had an average productivity of 3.24 t/ha, which is far below potential yields.[13]

The Bangladeshi government continues to invest in increasing the productivity of crop, fish, and livestock farmers through a large network of agricultural extension providers. Under normal circumstances, local agents engage in demonstrations and outreach, using an approach based on the T&V model discussed earlier.[14] IAPP provides an evolution of this strategy through the use of the farmer field school (FFS) approach.

In theory, Farmer field Schools should provide improved results over T&V through a number of mechanisms. One way to codify this is through Rogers' influential description of the "innovation-decision" process.[15] Rogers breaks down this decision into five steps: knowledge, persuasion, decision, implementation, and confirmation. While both FFS and T&V spread knowledge, FFS goes much further along each step of the process. By providing continued classes extolling the virtues of new technology a FFS is much more *persuasive* than the T&V approach. By providing input subsidies, FFS makes the initial adoption *decision* much easier. Through continued classes and instruction, the farmer is more likely to have a successful *implementation* of the technology, making him more likely to *confirm* his decision by continuing use of the technology. While there are many different approaches to FFS, they all provides a far more intensive and prolonged effort to spur technology adoption.

Farmer field schools were first introduced in Indonesia in the 1980s to promote integrated pest management (IPM), and spread to Bangladesh by 1994.[16] While they have subsequently been

---

[11] A. H. M. M. Haque, F. A. Elazegui, M. A. Taher Mia, M. M. Kamal and M. Manjurul Haque. "Increase in rice yield through the use of quality seeds in Bangladesh." African Journal of Agricultural Research Vol. 7(26), pp. 3819-3827, 10 July, 2012. http://www.academicjournals.org/ajar/PDF/pdf2012/10%20Jul/Haque%20et%20al.pdf

[12] Sayed Sarwer Hussain. "Bangladesh, Grain and Feed Annual 2012" USDA Foreigh Agricultural Service. http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Grain%20and%20Feed%20Annual_Dhaka_Bangladesh_2-22-2012.pdf

[13] Dey MM, Bose ML, Alam MF. 2008. Recommendation Domains for Pond Aquaculture. Country Case Study: Development and Status of Freshwater Aquaculture in Bangladesh. WorldFish Center Studies and Reviews No. 1872. The WorldFish Center, Penang, Malaysia. 73 p.

[14] Although T&V describes the traditional approach to agricultural extension in Bangladesh, it's not entirely accurate to assume that non-project areas (including control villages) are precisely practiving T&V. Some places might group-based approaches, but not with the level of organization and resources provided by IAPP.

[15] Rogers, Everett M. *Diffusion of innovations*. Simon and Schuster, 1995.

[16] Arnoud Braun, Janice Jiggins, Niels Röling, Henk van den Berg and Paul Snijders. "A Global Survey and Review of Farmer Field School Experience." Report prepared for the International Livestock Research Institute. June 12, 2006.

adopted to cover a wide variety of cropping patterns, much of the evidence on effectiveness of FFS covers IPM. A review of 25 impact evaluations of IPM FFS finds mostly positive effects, with FFS causing decreased pesticide usage and sometimes increased yields.[17] Studies on the impact of FFS for other technologies are far more limited. IFPRI conducted a study of FFS in three countries in East Africa, finding generally increased agricultural yields and farmer income for participants in FFS.[18] Additionally, DANIDA conducted an impact evaluation of two FFS programs in Bangladesh, finding a wide array of positive effects on farmer livelihoods.[19] However, both these evalautions relied on a retrospective creation of a control group using propensity score matching, calling into question the robustness of the results.

Despite a lack of hard evidence, the government of Bangladesh has been expanding the FFS approach past IPM[20]. However, FFS is more expensive than traditional extension approaches, so it is generally executed with donor support and is not yet widespread. This evaluation will assist the government of Bangladesh to understand the effectiveness of the intensive FFS approach versus traditional extension techniques, and to understand the extent to which it should be modified or scaled up.

4.2 Evaluation Questions

The main evaluation questions will be as follows:

- To what extent does the FFS approach promoted by IAPP cause increased and sustained technology adoption?
- What level of adoption is driven from increased subsidies (demonstration farmers) versus knowledge and learning (adoption farmers)?
- What are the differential effects for male versus female group members?
- Do the groups have spillover effects on other farmers who are not members?
- What are the long versus short run effects of IAPP? Do income effects allow continued adoption of improved crops?

[17] Henk van den Berg. "A Synthesis of 25 Impact Evaluations." Wageningen University, January 2004 Prepared for the Global IPM Facility.

[18] Kristin Davis, Ephraim Nkonya, Edward Kato, Daniel Ayalew Mekonnen, Martins Odendo, Richard Miiro, Jackson Nkuba. "Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa." IFPRI Discussion Paper 00992. June 2010.

[19] "Evaluation of the Farmer Field School Approach in the Agriculture Sector Programma Support Phase II, Bangladesh". Minsitry of Foreign Affairs of Denmark. November 2011.

[20] Hein W.L. Bijlmakers , Muhammad Ashraful Islam, "Changing the strategies of Farmer Field Schools in Bangladesh," Agricultures Nework. http://www.agriculturesnetwork.org/magazines/global/ecological-pest-management/changing-the-strategies-of-farmer-field-schools-in/

The Overall Project Evaluation is designed to most accurately pinpoint the effects of the crop and fisheries components, but will also produce results on livestock and water management.[21]

4.3 Evaluation Design

The evaluation is a randomized controlled trial, using a randomized phase-in of project villages for identification. The evaluation is designed to test both the long-term and short-term effects of the program.

For the Overall Project Evaluation, we include 96 villages that will receive at least crop and fisheries groups. Out of the 96 villages included in the evaluation, 48 will receive the project in 2012 (Treatment Villages), 24 will receive it in 2014 (Control Villages), and 24 (Long Term Control) will receive it in the last year of the project. These villages were randomly selected from the list of 96 villages that were eligible to begin the treatment in 2012. The villages that enter the project in later years will serve as control villages for those that enter the project in earlier years.

The randomization was jointly conducted by DIME and the PIU, using the random number generator in Microsoft Excel.

4.4 Sampling

4.4.1 Village Sampling

For the Overall Project Evaluation, 96 villages were sampled from 6 districts. In these districts, the selection of villages was conducted as follows:

1. We started with a list of all villages from each district that were eligible to begin project activities in 2012, and are slated to receive crop, fisheries, and livestock groups.
2. This list contained two villages per union (a smaller administrative unit). Therefore we randomly picked 8 unions per district to get our sampling frame of 96 villages.
3. Within each selected union, one village was randomly chosen as treatment and one as control.
4. At a later date, control villages will be divided into long term and short term control. Within each district, half of the unions will be randomly selected. These selected unions will have their control village assigned to be a long-term control (24 in total).

This results in a sample of 48 villages that will receive the project in 2012 (Treatment Group), 24 that will receive the project in 2014 (Short Term Control), and 24 that will receive the project in the final year of program operations (Long Term Control).

[21] The sampling has been done to pinpoint people who are eligible for fisheries and livestock groups, as these groups have the most restrictive criteria. These people are generally eligible for livestock and water management activities as well, and will receive them if it is logistically possible from a project perspective.

4.4.2 Individual Sampling

The project will primarily affect households who join farmers groups. However, these groups do not exist at baseline, and activities start quickly after group formation, leaving no time for a baseline to take place on these groups.[22] Hence we will sample based on a pre-identification of households that would be most likely to join farmer and fisheries groups, based on observable characteristics, and then re-sample to add other respondents during the midline and end line if needed. The sample was drawn from village-level lists of households that met the minimum eligibility criteria according to a rapid census, and also live close to one another, as geography is an important consideration in group formation. The IAPP team took this information into account when forming groups, which eased the operations on the ground and implies that most of our sampled population is likely to be included in the formed groups.[23]

Prior to sampling, a census was conducted in each village to determine group eligibility criteria and also determine other factors which make farmers more likely to join groups. Farmer groups consist of both demonstration farmers and adoption (non-demonstration) farmers, and there are different eligibility criteria for each. As we want to measure the program's impact on both types of farmers, our approach sampled a mix of demonstration and adoption farmers. The eligibility criteria were:

- **Crop, Adoption Farmer:** Small/Marginal farmer (< 2.5 ha cultivable land, prioritizing those with .21-5ha)
- **Crop, Demo Farmer:** Access to .5-1 ha of cultivatable land, willingness to engage in demonstration
- **Fisheries, Adoption Farmer**: Small/Marginal farmer, owns or rents a fish ponds between 13-100 decimals
- **Fisheries, Demo Farmer:** Small/Marginal farmer, owns or rents a fish ponds between 13-50 decimals

The other factors considered were:

- Size of family (larger families likely to join)
- Whether Crops/Fisheries are the primary source of income
- Grow grains/legumes/oilseeds (more likely to join crop group)
- Willingness to engage in demonstration

Based on these criteria, we developed a list of people most likely to join farming or fisheries groups, and then selected a cluster of 8 people (each for fish and crops) to sample for the baseline.[24]

[22] We did survey after group formation for the Demonstration Plot Evaluation, but due to limited time it was not possible to survey the villages for the Overall Project Evaluation as well in this short time frame.
[23] The groups are formed using a village participatory approach, so the IAPP team has limited control over who ends up in groups.
[24] The selection algorithm used GIS data select based on eligibility and geographical distance from a central node.

Additionally, one of the evaluation questions is to understand the spillover of project effects to people who do not join groups. This strategy will result in sampling some people who are eligible to join groups but do not end up doing so, which will allow us to test for spillover effects of the program.

4.4.3 Power Calculations

As in the previous section, power calculations were undertaken primarily using data from a recent USAID agricultural survey conducted by IFPRI in Barisal district (which is one of the IAPP districts). Estimates of the predictive power of baseline statistics come from a panel dataset in India.[25]

The IFPRI survey contains data on three of our primary outcome indicators: paddy yield, fish yield and household expenditure. Therefore, we concentrate on these measures for the power calculations. For an estimate of the program's effects, we turn to the results framework, which predicts a short-term increase of 300 kg/ha for paddy and 500 kg/ha for fish. It predicts a long-term increase of 500 kg/ha for paddy and 700 kg/ha for fish.

Although not included in the results framework, we also seek to measure the effect on household income, which is proxied by per capita expenditure. The IAPP project documents do not contain an estimate for the expected effect on per-capita income, so we apply a standard assumption of a standard effect size of .2 for the short term, and .25 in the long term.

Finally, we need to account for the increase in precision provided by our baseline. As the IFPRI survey is a cross-section, it is not informative about the expected R^2 from baseline data. Therefore, we use estimates from an agricultural survey conducted in India to make informed guesses about the R^2 for income and yield. This survey did not contain data on fish yields, so we estimate the R^2 for fish to be zero in order to remain conservative.

The data necessary for the power calculations is illustrated in the table below:

Table 2: Power Calculation Data

	Average	Standard Dev	Incremental Change Expected (Long Term)	Expected Normalized Effect Size (Long Term)	Incremental Change Expected (Short Term)	Expected Normalized Effect Size (Short Term)	Village Intracluster Correlation (ICC)	R^2 from baseline
Paddy Yield (kg/ha)	3410	1977	500	.25	300	.15	.146	.5

[25] For the R^2 We use agricultural revenues as a proxy for yield and total expenditure to proxy for income.

Fish Yield (kg/ha)	1836	1944	700	.36	500	.25	.054	0
Expenditure (taka/month /hh)	2713	1553	311	.25	389	.2	.098	.25

For crops, in addition to the 96 villages sampled for the Overall Project Evaluation, we will also use data from 110 villages that are part of the Demonstration Plot Evaluation to increase precision.

In some specifications, the number of villages is not balanced among treatment and control or the number of people sampled is not the same in all villages. When this lack of balance occurs, we perform power calculations using formulas assuming balance, but replacing the unbalanced elements with their harmonic means.[26] For the short term analysis of crops, our sample has 206 villages with an equivalent of 10 people surveyed per village. For long term analysis of crops, we have an equivalent of 129 villages (again with 10 people surveyed).

Samples and their predicted MDES are summarized in Table 2:

Table 3: Sampling Results (85% power)

	Individuals / group (Balanced Equivalent)	Number of groups (Short Term)	Expected Normalized Effect Size (Short Term)	Short Term MDES	Number of groups (Long Term, Balanced Equivalent)	Expected Normalized Effect Size (Long Term)	Long Term MDES
Paddy Yield (kg/ha)	10	206	.15	.17	129	.25	.21
Fish Yield (kg/ha)	8	96	.25	.27	72	.36	.30
Expenditure (taka/month /person)	15	412	.2	.12	227	.25	.16

[26] This strategy is suggested in the user's guide for Optimal Design software, which was used to perform these calculations. User guide can be found here: http://pikachu.harvard.edu/od/od-manual-20111016-v300.pdf.

For crops, at a power of 85% these numbers result in a minimum detectable effect size (MDES) of .17 for the short term comparison.[27] For the long term comparison, we foresee a MDES of .21.

For fisheries, we will have 96 villages in the short term and the equivalent of 72 in the long term. Based on surveying 8 people per group, this leads to an MDES of .27 in the short run, and .30 in the long run.

We can calculate income changes using people from all crop and fisheries groups. If we pool the income effects from both components, and assume clustering at the group level, we end up with 412 groups short term and 227 long term, with a sample of 10 people per group. This achieves an MDES of .15 in the short term and .18 in the long term.

As shown in Table 2, nearly all the results achieve comfortable power with the given sample. The one exception is the short term estimate of paddy yield, where the expected normalized effect size is a bit smaller than the short term MDES. In this case, the power to detect the predicted effect size of .15 is 79%, which is a bit lower than ideal. However due to logistical constraints it was not feasible to increase the sample size.

5. Data

The data for the impact evaluation (both the Overall Project Evaluation and the Demonstration Plot Evaluation) will come from three household surveys: a baseline survey conducted from Aug-Oct 2012, a follow-up to be conducted from July-Sept 2014, and an endline to be conducted from July-Sept 2016. The survey questionnaire will be guided by project goals, the results framework, and the GAFSP monitoring and evaluation framework. It will collect detailed data on crop, fisheries, and livestock practices, and will directly report on the causal impacts of the project on the Project Development Objectives as defined in the results framework:

- *Improved Crop Productivity:* The main indicator will be crop yields, measured in kg/hectare, during all three growing seasons (particularly Kharif and Rabi). We will also the exact variety of crop grown, and usage of inputs such as irrigation, improved seed, fertilizer, farmyard manure, and green manure, along with adoption of promoted crops such as mung and cucumber. The collection of inputs will allow measurement of not both revenue and agricultural profits.
- *Improved Fisheries Productivity:* The main indicator will be fish yields, measured by kg/hectare. We will also track the adoption of improved fish varieties and cultivation practices.
- *Improved Livestock Productivity:* For cows, we will measure liters of milk per cow per day. For goats and chickens, we will look at sales and consumption of meat and eggs.

[27] All calculations assume a chance of Type I error to be 5%.

We will also include the following indicators from the GAFSP monitoring and evaluation framework:

- Total Income
- Total income from agriculture (disaggregated by crop, fisheries, and livestock)
- Food security (measured by Household Hunger Score and Dietary Diversity Score)

Since many of these new technologies will require increased spending on inputs (including family labor), we will also measure input spending closely in order to evaluate farm profits. We also will map the social networks within groups in order to understand how relationships affect technology adoption.

Costing exercises will also be carried out to understand how much it would cost to scale up both the overall IAPP FFS model and the different demonstration approaches in the Demonstration Plot Evaluation. This will allow us to conduct cost-benefit analysis.

6. Internal and External Validity

The main threat to internal validity of this study is that the assignment of treatment and control villages may not be respected. Although the PIU has agreed to the evaluation plan, local officials are not always aware of the need to avoid operating in control groups, and exactly where and how to implement the different interventions in the Demonstration Plot Evaluation.

In order to mitigate this risk, DIME is working with Innovations for Poverty Action (IPA) to closely monitor the situation on the ground, and work closely with local officials to ensure that they respect the evaluation plan to the best of their abilities. During the implementation of the Demonstration Plot Evaluation, monitors from IPA will attend the group meetings to ensure that the interventions are taking place in the correct villages, and to gather data on who agrees to be a demonstration farmer. Similarly, IPA monitors will assist with the measurement of crops which determine the incentive payout.

The main threat to external validity is if the villages selected for our sample are not representative of Bangladesh as a whole. As we were provided with a list of eligible villages from the district officials, it is not clear whether they constitute a representative sample. However, it was necessary for the evaluation to work within the constraints of the project, which could only work in areas where it had the required resources in place.

6. Evaluation Team

This evaluation will be managed by DIME, with close collaboration with the IAPP World Bank Project Team and the IAPP PIU. The DIME team consists of:

- Florence Kondylis, Economist: Task Team Leader, GAFSP-DIME Impact Evaluation Portfolio
- Daniel Stein, Economist: GAFSP-DIME Impact Evaluation Portfolio
- Maria Jones, Research Analyst: GAFSP-DIME Impact Evaluation Portfolio

- Cindy Sobieski, Field Coordinator: IAPP Impact Evaluation Field Coordinator

DIME is also working with Innovations for Poverty Action, a research NGO that specializes in impact evaluations. IPA will provide research support on evaluation design, data gathering, and implementation monitoring. The IPA team is led by Mushfiq Mobarak, an associate professor at the Yale University School of Management.

7. Budget

Based upon quotes from our selected data collection/research firm (Innovations for Poverty Action) and our prior ample experience with impact evaluations, we expect the following costs. The costs for the survey are as follows:

1. Census to create sample: $33,560.
2. Baseline Survey: $166,003
3. Monitoring and Oversight of Demo Plot Evaluation: $48,708
4. Midline Survey: $181,248
5. Endline Survey:$190,000

There will also be costs associated with oversight of the evaluation, and development of reports

1. Cost of a field coordinator for four years at $40,000/yr = $160,000
2. Total travel costs for lead researchers during 5-year period = $75,000
3. Research assistance for data cleaning and report preparation: $37,500/yr, total of 1 year = $37,500.

We estimate the total cost to be $891,389. These funds have been made available by GAFSP, and will not come out of project funds.

8. Timeline

The first steps of the project proceeded as follows:

1. **Identification of project villages**: Each department involved in the project (DAE, DoF, DLS, BADC), selects the villages in which they are able to work. The PIU will then select the villages for inclusion in the project, giving preference to villages where all project components are feasible. Working with local officials, the PIU successfully completed this on August 1, 2012.

2. **Selection of IE Village Sample**: As the project hopes to cover over 5,000 villages, the impact evaluation will necessarily need to concentrate on a subset (316) of these villages. The subset conforms to the sampling strategy described above. This process was completed by August 31, 2012.

3. **Census of Households for Surveying**: In order to conduct sampling for the Overall Project Evaluation, a brief census of all households in sampled villages was conducted. The census covered eligibility for joining different project components: land ownership, livestock ownership, and access to fish ponds. This was completed by Sept 20, 2012.

4. **Sample of Households to Be Surveyed:** Using the census, households were chosen to be part of the baseline survey for the Technology Adoption Evaluation. For the demo plot evaluation, households were randomly sampled after farmer group formation. This sampling was completed during Sept-Oct 2012.

5. **Randomized Selection of Treatment Arms and Control Villages:** Once the impact evaluation sample was selected, the sample was randomly divided into treatment and control arms, as described in the evaluation strategy above. This was completed by October 1, 2012.

6. **Baseline Survey**: A baseline survey of sampled households in both control and treatment villages was completed by October 31, 2012.

7. **Implementation of Demonstration Plot Evaluation:** DIME is monitoring and collecting data on selection of demonstration farmers for those groups that are part of the demonstration plot evaluation. This will be completed by December 15, 2012.

The long-term timeline can be found in Figure 3 below:

Figure 3: Long Term Timeline