R E P O R T

MCC Ghana Impact Evaluation Services

Evaluation Design Report -- UPDATE

JUNE 12, 2013

REVISED ON
SEPTEMBER 8, 2017

SUBMITTED TO:
Millennium Challenge Corporation
Contracts and Grants Management Division
Attention: William Hummel, Senior Contract Specialist
875 15th Street, NW
Washington, DC 20005
Telephone: (202) 521-3600

SUBMITTED BY:
NORC at the University of Chicago
55 East Monroe Street
30th Floor
Chicago, IL 60603
Telephone: (312) 759-4000
Fax: (312) 759-4004

NORC at the UNIVERSITY of CHICAGO

TABLE OF CONTENTS

A. INTRODUCTION

In 2011, MCC contracted NORC to assess the impact of up to five activities under the MCC Program in Ghana using the most rigorous methods possible. These activities included:

1. Agriculture Project, Post-Harvest Activity and Community Services Project, Electrification Sub-Activity
2. Agriculture Project, Irrigation Activity
3. Agriculture Project, Credit Activity
4. Rural Development Project, Community Services Activity, Education Sub-Activity
5. Rural Development Project, Financial Services

In December 2011, NORC submitted to MCC a Design Report that outlined proposed evaluation designs and potential data sources for the six activities/sub-activities. These designs were based on information gathered through the following activities:

- Review of background information for each of the six activities, including objectives, implementation details, background review of relevant geographical areas and economic sectors, as well as some baseline data that was available.

- Meetings with relevant staff from MCC, who provided additional details about the six interventions.

- A two-week trip to Ghana by key members of the NORC team in November 2011 to meet with MCC, MiDA, supervising and implementing agencies, and other stakeholders to gather critical information on project implementation and coverage, and data to inform decisions regarding evaluation designs.

Since the submission of the Design Report in December 2011, there have been adjustments to some of the evaluation designs that were originally envisioned for some of the activities, while others have remained unchanged. For other activities, MCC has opted not to pursue an impact evaluation. Below we present updated evaluation designs for the interventions that are currently being evaluated.

B. UPDATED EVALUATION DESIGNS

In this section, we discuss each of the MiDA activities and sub-activities separately. For each activity/sub-activity that we are conducting an evaluation for, we present the following:

- Background information about the intervention
- Key evaluation hypotheses and impact indicators
- Data sources
- Evaluation Methodologies, associated risk factors
- Key considerations

B.1 Agriculture Project: Post Harvest Activity

The Post-Harvest Activity aims to provide infrastructure – coolers, pack houses, agribusiness centers, and a Perishable Cargo Centre at the Accra airport – to producers and other actors in the value chains for horticultural exports and grains for local markets. These facilities would help maintain the quality of these agricultural products as they are transported from the farms to their markets.

Post-harvest infrastructure, supported by adequate power sources, can affect the incomes of farmers, exporters, and other actors in the respective value chains by providing storage for farm products; reducing post-harvest losses of fruits, vegetables, and grain crops; and improve their quality at the time of arrival at their markets. These investments can thereby increase market prices and/or open or ease access into new markets (export markets, for example) for these agricultural products.

B1.1 Agribusiness Centres

Background Information

The Millennium Development Authority (MiDA) constructed ten Agribusiness Centers (ABCs) that provide services for the initial processing, storage, and marketing of grain crops produced by farmer-based organizations (FBOs) within their respective intervention areas. Each ABC was outfitted with specialized equipment for processing either rice or maize, although its complement of installed equipment could later be modified for processing other grain crops, such as soybeans, as their operations expand. The initial crop selected for processing at the respective ABCs would be based on the prevalent crop that is grown in the area. Each ABC is designed to store approximately 1,000 tons of grain and will serve as a grain processing and marketing center for FBO members located in the vicinity of the center, within a radius of approximately 20 kilometers.

The ABCs are expected to provide for-fee grain processing services including maize shelling or paddy rice de-husking and de-stoning, along with grain drying, cleaning, sorting, selecting, bagging, palletizing, and storing. If desired, the ABCs will market the grain inventory stored on behalf of its FBO clients; otherwise, the ABC will store the grain securely until it is sold directly by the client. In addition to grain processing, storage, and marketing services, the ABCs will sell inputs to its FBO members, including improved seed, fertilizer, and farm chemicals. It will also provide tractor services to small farmers who are members of its affiliated FBOs to help them prepare their land for planting.

The ABCs are to be privately-owned, privately-operated, profit-making service organizations. Each facility will be jointly owned by a private entrepreneur, known as a “lead investor”, along with twenty farmer-based organizations, each of which has a membership of approximately 50 members.[1] Each ABC is to be managed by the respective lead investor, selected by MiDA through a competitive bidding process, to own 70 percent of the ABC.

[1] Note that more than one member of a farmer household may belong to an FBO. This means that an FBO member does not necessarily represent one farm or one rural household.

The ABCs are generally located within those areas where there are heavy concentrations of grain production. The lead investors have been instrumental in deciding the specific location of the respective ABC within the grain producing areas, and most are located along main roads with a nearby supply of electric power and water.

All 10 ABCs were constructed, equipped and transferred to the new private sector companies (made up from private sector investors and FBOs) by the end of the Compact period. MCC reports that these ABCs are fully functional at present.

MiDA delegated the process of selecting the FBO partners at each ABC to its regional implementation consultants (RICs). Once the location of the ABC was decided, the RIC drew an imaginary circle with a 20-kilometer radius around the location, and all the FBOs within the defined area that had been registered and trained under the MiDA program became candidates for membership in the ABC. The RIC then invited these FBO leaders to an executive session to discuss the possibility of creating a FBO Union that would participate in the proposed ABC. Through this orientation process the FBO leaders became sensitized to the potential benefits of cultivating the selected crops and marketing them through the ABCs. Additional meetings and follow-up resulted in the creation of an FBO Union composed of 20 FBOs with approximately 1000 members that would become co-owner of the respective ABC. After the initial FBO selection had taken place, there was a second round of fine-tuning for the participants to reach the final composition of the FBO Union. Some FBO members decided not to participate in the venture, and in a few instances, FBOs with extremely poor credit repayment records were rejected. The rejection was conducted on a case-by-case basis, based on a subjective assessment by the MiDA staff of the FBO's past credit payment performance.

Evaluation Hypotheses and Impact Indicators

Like other MIDA post-harvest infrastructure investments, ABCs aim to increase the livelihood of small farmers. This intervention in particular was designed to resolve three major issues that affect commercial agricultural production by small farmers: 1) deficient handling, processing and storage of agricultural products after harvest, 2) weak marketing systems that results in below-market prices for commodities produced by small farmers, and 3) lack of reliable agricultural input supplies and farm equipment services commercially available to small farmers.

The hypothesis is that solving the above problems should translate into 1) a reduction of post-harvest losses, 2) better market prices, and 3) higher crop yields. All this in turn increases total production, total profits, and therefore farmers' income from crop harvest.

There are several indicators that can be used to measure the impact of the ABCs on the wellbeing and productivity of small farmers. They include the following:

1. Total annual household income.
2. Total annual farmer revenue from maize and rice production
3. Annual sales volume (in kilograms, for example) of the targeted crops per household.
4. Percentage post-harvest losses for the targeted crops
5. Crop yield – net amount of grain produced per unit area (e.g., kilograms of maize or rice produced per hectare)

Additionally, other intermediate outcomes such as use of fertilizers, insecticide, fungicide, and other production inputs and unit prices for the crops, can be measured in order to understand intermediate impacts.

Without additional assumptions, for which there is no data-based evidence, there is no reason to think that this intervention should have differential effects by sex or age of the farmer. Despite this, we can try estimating effects by gender.

Evaluation Methodology

As stated in NORC's December 2011 Design Report, we propose to use a double-difference estimator with matching or covariate controls as appropriate, or a combined regression analysis/matching approach to assess changes in farmers income, production, crop revenue and post-harvest losses related to access to ABCs. Alternative estimations will be provided. Under this approach, find units in the comparison group that are as similar as possible to the treated units by computing the probability that a unit will belong to the treatment based on its observable characteristics. The goal is to mimic a randomized assignment when it does not exist. We will use available data from the FBO Survey and the 2010 Census, as well as GIS data on climate and topography for the matching process.

Our sample consists of farmers that belong to FBOs who will have access to and will benefit from the ABCs, and a comparison group comprised of farmers in similar FBOs that will not benefit from the ABCs. The treatment group for each ABC will be comprised of farmers interviewed in the baseline survey that belong to FBOs within an approximately 20-kilometer radius around the ABC. The purpose of including all FBOs within the 20 km radius in the treatment group (as opposed to only the 20 co-owners of the respective ABC) is twofold: 1) we want to avoid selection into treatment bias (see details of selection process in Background Information Section above), and 2) we want to include all farmers that might benefit from the ABCs facilities even if they are not co-owners.

In this methodology, it is also essential to identify an appropriate comparison group. In conversations the NORC team had with local experts during the design trip to Ghana, the MIDA team and other stakeholders suggested that the comparison group for each ABC can be selected from MiDA-trained FBOs located outside the approximately 20-kilometer radius around the ABC. In this regard, those FBOs forming the comparison group would be located between two concentric circles, centered on the respective ABC. The inner circle will have a radius of approximately 20 kilometers from its center, corresponding to the ABC location, and the outer circle will have a radius of approximately 30 kilometers from the ABC. Farmers linked to the MiDA-trained FBOs that were included in the baseline survey and are located in the area between these two circles would be defined as the comparison group.

In addition to the difference-in-difference methodology, we propose to use an Instrumental Variable two stage least square (IV-2SLS) approach, using a model where "treatment" is instrumented in a first stage by "distance from farmer to closest ABC". We will still select all sampled farmers within a radius of 30 km but rather than considering everyone within 20km as “treated” and the rest "controls," we will let the variable "distance to ABC" instrument participation into treatment. A second stage will estimate the effect of instrumented participation on outcomes of interest.

Data Sources

For this evaluation, we use GIS data compiled by MiDA to geolocate ABCs and affiliated FBOs for the purpose of identifying treatment and comparison FBOs and farmers based on their distance from the ABC.

The main source of data for this evaluation comes from primary data collection. In 2012, NORC conducted a baseline survey of a sample of 2012 treatment and comparison farmers, as defined above, based on the FBO data lists provided by MIDA. We collected data on household and farmer characteristics, crop and yield information, plot measures, and a host of other indicators. The endline survey will be conducted among the same farmers, creating a pane dataset. The post-intervention data collection could be carried out at any time after approximately two years from the time that each FBO began operating, or at minimum, after two crop cycles of the targeted crops have been completed. The exact timing of the endline data collection is yet to be determined.

Key Considerations

It is important to keep in mind when interpreting the effects of the ABCs that this sub-activity involves more aspects than simply the physical ABC and associated services. As described in Section A1.1 above, the "lead investors" were selected in a competitive bidding process and they are important in the management of the ABCs. This is a central factor that must be taken into consideration when replicating this type of intervention. The location of the ABC is also not random. The Centers are located near main roads, power sources and water. Our analysis should evaluate the effects of this "ABC package" on farmers' economic wellbeing, and the results should be interpreted accordingly.

B1.2 Public Pack houses – No Evaluation

Background Information

MiDA's portfolio of post-harvest activities included support for the construction and equipping of three large, state-of-the-art public pack houses (PPH) in Ghana's Southwestern Horticultural Belt. Two of the pack houses will serve the export pineapple agro-industry, and the third will be used for mango exports. The two pineapple facilities are located in the districts of Gomoa and Akwapin South and the pack house for mangos is in the Yilo Krobo district. The three pack houses have automated packing lines with the capacity for moving freshly harvested fruit on a conveyor to stations where it is automatically separated and sorted into batches of fruit of similar size that are then manually packed into carton boxes containing a standard weight. The packed fruit boxes are palletized for ease of handling and, once the palletized unit has been quick-cooled, it is stored in refrigerated rooms at the pack house until it is loaded into refrigerated containers and transported to the Tema Port. At Tema, the fresh fruit remains in the sealed container until the scheduled arrival of a container ship that transports the container with its refrigerated cargo to European ports for discharge and distribution to fruit wholesalers, brokers, and supermarket chains. The European market has extremely demanding quality standards for fresh fruit, and imposes rigorous procedures for post-harvest handling and temperature control as essential elements of an export program that serves this market. Consequently, MiDA's state-of-the-art pack houses and the capability to support the first link of the "cold chain" for export horticultural products from the farm to the final customer are an important part of Ghana's strategy to increase its horticulture exports to markets overseas. The public pack houses will each serve from 500-600 farmers, and these will considerably increase the capacity for pineapple exports from nearby communities. The public pack houses are privately owned and operated. The co-owners of the pineapple pack houses aim to be the anchor farmer (60 percent ownership) and the affiliated FBOs (40 percent ownership) that serve as pineapple outgrowers (contract farmers) to the anchor farmer. In the case of the public pack house for mangos, where no anchor farmer is available, its owners will be the Dangme Union and its FBO members.

Evaluation Hypotheses and Impact Indicators

The public pack house will provide packing services for agricultural products to the anchor farmer, the associated FBOs and their members, as well as other independent FBOs that use the facility. This will have several positive effects:

- A well-managed, functional pack house should improve the quality of the agricultural products that are offered for sale in the targeted market reducing losses from handling and transportation and claims from buyers for inferior product quality.
- A pack house that consistently ships good quality agricultural products can enable the exporter to increase his or her share of market for that product by displacing inferior quality products. This could increase product sales.
- Consistent quality can improve the position of the exporter and the average market prices received.

The idea is that the effect of improved product quality and consistent access to markets, in particular international markets, will provide higher average selling prices. Better prices are expected to increase the amount of income received by small farmers. To the extent that part of the increased income is used for orchard maintenance and reinvestment, the amount of mangos produced will gradually expand income.

Evaluation Methodology Proposed in Evaluation Design

Originally, for this intervention, NORC proposed an evaluation using the same approach as that proposed for the evaluation of ABCs. However the number of farmers that produce pineapple and mangos represented in the MIDA FBO survey was discovered to be quite small. This is particularly true for mango producers. Although we did not have final numbers for mango and pineapple farmers in the baseline MIDA FBO Survey due to some missing data, we estimated them at approximately 45 treatment and 45 comparison mango producers and 95 treatment and 95 comparison pineapple farmers.

Estimates of impact based on small samples tend to be imprecise and in such cases detecting the impact of an intervention or program can be difficult. The power calculations presented in Annex 2, show that for the case of the Mango PPH, with only 45 farmers in the treatment group and 45 in the control group, the effect size would have to be 0.55 or larger to be able to detect an income change with a moderately high power of 0.8. A 55 percent average change in income is very ambitious, and in the event that it does not occur, we run the risk of concluding that the intervention did not have a positive effect.

In the case of pineapple farmers, the sample is likely to be around 190, with 95 treated farmers and 95 controls. In this case, the required change in income for detecting a change at a reasonable level of power would be lower than 55 percent, but still quite high.

After lengthy discussion on this issue, MCC and NORC agreed that this activity would not be evaluated.

B1.3 SPEG Loans for Cold Rooms and Packing Lines – No Evaluation

This intervention is similar to the PPH activity in terms of the infrastructure it provides. In 2008, MiDA provided a conditional grant in the amount of US $5.3 million to the Sea-Freight Pineapple Exporters of Ghana (SPEG) to create a loan program administered by SPEG that would enable its members to construct cold rooms, install automated packing lines and to provide a stand-by generator for their pack houses. In those cases where the electric power grid did not reach the pack house, as was the case at the 2K farm, MiDA's Rural Development Project provided electricity to the pack houses and to nearby communities by constructing an electric power line.

In September 2008, MiDA announced the approval of the first loan tranche in the amount of US $2.17 million for seven SPEG exporters who made up the first phase of the loan program. The loans were provided for a five-year term at a flat interest rate of 5 percent per year, with annual loan repayments amounting to 20 percent of the original loan amount, plus interest. Unfortunately, this first group of loan recipients was not able to repay their loans as scheduled. As of September 2011 around 17% of the loan has been repaid, and SPEG has unilaterally re-scheduled its repayment of the MiDA debt (without obtaining formal agreement from MiDA), and plans to complete all loan payments under its revised schedule by 2015.

Hypotheses and Impact Indicators

The competitiveness of Ghana's export pineapple industry depends on the quality of the fruit it exports to international markets and on the cost of placing its export pineapples in these markets. Ghana cannot be competitive in export markets without packing facilities that minimize post-harvest handling damage and that enhance the capability of the exporter to provide good, consistent quality of the exported fruit under efficient, low-cost packing methods. Another key element in post-harvest handling of fresh pineapples is the ability to quickly reduce the temperature of the fruit ("remove the field heat") to the optimum storage temperature, and to maintain that temperature until the fruit is delivered to the buyer. Both these elements are required to compete in Ghana's export markets. The loans provided to the seven SPEG exporters were designed to increase their ability to compete in European markets.

The effect of these improvements on the SPEG exporters should follow a model very similar to the one described for the public pack houses. However, it may differ in the expected impact on small farmers. In light of the information that MiDA provided the evaluation team, it appears that the SPEG equipment loans provide the greatest benefit to the exporters, whereas their impact on small farmers should be quite limited given that only three of the seven exporters that received equipment loans – Bomart, Prudent, and Georgefields – have outgrower programs. Each of these exporters is associated with a single FBO that has about a dozen members as outgrowers. The remaining four exporters - Jei River Farms, Koranco Farms, Gold Coast Fruits and 2K Farms - do not have an outgrower program.

While the evaluation model described for PPHs is relevant for the cases where exporters work with outgrowers, the rest of the exporters are unlikely to generate positive spillover effects for small farmers. The main reason is because in general, outgrower programs by most SPEG exporters are extremely limited, and without outgrower programs, small farmers receive little benefit from these investments. If these investments have any effect on small farmers, it is more probable that such an effect would come from increases in paid employment at the exporter farms and pack houses.

Evaluation Design

Originally NORC proposed a design to evaluate the effects on small farmers linked to exporters that work with outgrowers, using the same methodology suggested for the PPH and ABC evaluations. However, as we mention above, there are only three such exporters, and each of them is associated with only one FBO. This means that we will on have a very small group of treated farmers that were covered in the FBO Survey (15 at best) and it is highly unlikely that under such conditions even a very large impact could be statistically detected.

As it was the case of the PPH, MCC discussed the issue with NORC and decide not to evaluate this activity.

B1.4 Perishable Cargo Center

Background Information

MiDA funded the construction and equipping of a US $2.7 million perishable cargo center at the Kotoka International Airport (KIA) in Accra under its Agricultural Project, to support increased exports of fresh fruit and vegetables from Ghana.

In 2011, KIA exported around 20,000 tons of fresh agricultural products annually. Just about all vegetable and cassava exports were shipped by air to overseas markets; around 80 percent of fresh papaya exports are shipped as air cargo, and around 10 percent of pineapple exports - mostly pre-cut packaged pineapple chunks exported by Blue Skies, Ltd. – are shipped by air from KIA.

The perishable cargo center (PCC) is the final link of an integrated cold chain for the horticultural sub-sector that begins at the pack houses where products are initially cooled.

At present there is no packing shed at KIA where fresh produce can be consolidated, nor is there a cold storage facility to maintain the quality of exported fresh products. Even when the cargo arrives by refrigerated truck, the shipment must be discharged, palletized, and held for export at ambient temperature, breaking the "cold chain."

The PCC will be owned by the Ghana Airports Company Limited (GACL) and will be managed and operated by a consortium headed by Ghana Air under a concession from GACL. The consortium was selected through a public bidding process.

The PCC will be a public facility, providing for-fee export services to all horticultural exporters in Ghana, and possibly even some in neighboring countries, who wish to use its services. Its clients will be drawn from the community of active exporters.

Hypotheses and Impact Indicators

The PCC is a key element in the achievement of better quality of exported fresh produce shipped by air to markets overseas. The PCC could lead to higher market prices and increased volumes for its normal fruit and vegetable exports, as well as stimulating export growth in emerging export products such as cut flowers.

The benefits to be derived from the PCC are summarized as follows:

1. The perishable cargo center will serve Ghana's fresh horticulture exporters by providing the range of services required to ship their fresh products by air to buyers in foreign markets. Of critical importance will be availability of refrigerated storage at the PCC, which will keep the export products in good, fresh condition to maintain quality and shelf life. A second benefit is that with cold storage, the exporters will be able to deliver their products ahead of time to the airport, and will not have to precisely schedule their deliveries around the aircraft departure time. By ensuring the quality of export products, the PCC will enable horticultural exporters to effectively compete in upscale markets for fresh fruit and vegetables in the European Union (EU). The lack of a modern facility for handling fresh horticultural products has generally limited exports from Ghana to the lower tier of EU markets due to buyers' concerns of poor product quality. While data are not readily available on the amount of product loss due to spoilage under current handling conditions, a knowledgeable estimate would be a loss of around 5 percent. The greatest loss of exporter income under current conditions results from quality claims, and the foregone revenue resulting from selling into downscale markets.

2. Increased competitiveness of fresh fruit and vegetables from Ghana sold into European markets will result in an increased amount of products exported, with higher prices relative to the average market price for the export product. Many exporters, particularly those who export fresh vegetables, rely on small farmers to achieve a critical mass of their export product. The PCC will spur the growth of the export supply chain, and the corresponding benefits will be shared by small farmers who supply these export products. However, this benefit would be extremely difficult to measure, given the number of steps the small farmers are removed from the intervention.

3. Improved handling and storage of export horticultural products will reduce the amount of claims for poor quality by foreign importers, against the exporters in Ghana.

4. The greater convenience and improved logistics for handling perishable products at the PCC will make it easier for exporters to comply with aircraft schedules and to meet product delivery deadlines. This will result in increased exports and fewer shipments held over.

Based on these hypotheses, we propose to use the following impact indicators:

1. Volume of exports
2. Claims from foreign exporters
3. Product loss volumes

Evaluation Methodology and Impact Indicators

As described earlier, the services provided by the PCC will be available to any exporter of horticultural products in Ghana, and it is likely that most fruit and vegetable exporters will use the facility. While in some cases groups of small farmers supply the community of exporters with agricultural export products, identifying these groups and determining the impact of the PCC on their farming operations would be extremely difficult, and impractical. As such, we do not propose to evaluate the impact of the PCC on small farmers.

Instead, we propose to conduct the evaluation using the exporter as the unit of analysis. We will focus our analysis on around 30 exporters of horticultural products who consistently air freight their products from KIA. Since most exporters will use the PCC when it becomes available, we see no opportunity to construct a comparison group of exporters who do not access the facility. Therefore, the evaluation approach will be a simple pre- and post-intervention design, which uses data from a variety of sources for a pre-PCC baseline period and a post-PCC endline period. This approach, in the absence of a control group, precludes the possibility of attributing change to the availability of the cargo center.

Data Sources

A broad spectrum of exporters will likely use the services of the PCC, including shippers of fresh fruit, vegetables, and root crops; as well as exporters of pre-cut, packaged fruit and vegetables. Consequently, there is no single organization that can serve as the source of data to measure the impact of the PCC. Instead, both baseline and endline data must be obtained from the Ghana Airports Company. In addition, we will explore the possibility of using data from those individual exporters who have sold abroad their products through KIA.

These sources would provide data on the following indicators: the annual throughput (metric tons) of fresh horticultural exports that pass through the perishable cargo center at KIA; the annual volume (kilograms) and value (US $) of fresh horticultural exports, by commodity, that are exported through the PCC; and the annual amount of claims against exporters as a percentage of the annual value of horticultural products that pass through the PCC.

B.2 Agriculture Project: Irrigation Activity

Background Information

MIDA funded the renovation of two irrigation schemes in the Tolon Kumbungu district in the Northern Agricultural Zone, and the construction of a new scheme in the North Tongu district in the Southeastern Horticultural Belt. The northern schemes are the Bontanga Irrigation Project and the Golinga Irrigation Project, both located near Tamale, the regional capital. The new southeastern scheme, known as the Kpong Left Bank Irrigation Scheme, is located in the Volta Region near the village of Torgorme.

The following table summarizes the most important characteristics of the three irrigation schemes[2]:

<table>
<tr><th colspan="6">MiDA Construction and Renovation of Irrigation Schemes</th></tr>
<tr><th>Scheme Name</th><th>Area Small Farmers (HA)</th><th>Area Anchor Farm (Ha.)</th><th>No. FBOs</th><th>No. Small Farmers</th><th>Cost (US $000)</th></tr>
<tr><td>Kpong Left Bank</td><td>450</td><td>1070</td><td>15</td><td>746</td><td>10,881</td></tr>
<tr><td>Bontanga</td><td>495</td><td>315</td><td>10</td><td>528</td><td rowspan="2">3,047</td></tr>
<tr><td>Golinga</td><td>40</td><td>None</td><td>5</td><td>246</td></tr>
<tr><td colspan="6">Source: MiDA and IFDC technical staff</td></tr>
</table>

The irrigation intervention was predicated on the expectation that:

- Small farmers operating within each of the three irrigation schemes will have the opportunity to participate in contract farming arrangements with a large, commercial farm known as an "anchor farm" located near the small farmer irrigation scheme. The anchor farmer would have access to irrigation water from the main canal that carries water from the reservoir to the irrigated area for small farmers and, in turn, would be required to pump irrigation water onto their farm, since the terrain does not permit gravity flow to these farms.
- The anchor farmers would provide training, technical assistance, and seed to the contracted small farmers through their FBO Unions, for the production of the required crops. In addition to providing market outlets for their designated crops, the anchor farmers would help its small farmers to comply with international standards for export products (eg, GlobalGap) as required.

Kpong Left Bank: The Kpong irrigation system is being constructed just outside MiDA's targeted intervention area of thirty districts. The anchor farmer at the scheme, Vegpro, was planning to farm an irrigated area of 1,070 hectares located adjacent to the small farmer irrigation scheme. *Construction of this irrigation scheme began on January 21, 2011 and per reports from MCC, was expected to be completed between June and September 2013.*

[2] The evaluation design team was informed by MCC that the costs shown in this table (provided by MiDA) include neither the feasibility studies nor construction supervision, nor the investments done for the anchor farmers in the irrigation schemes.

Bontanga: The Bontanga irrigation project is the largest irrigation scheme in the Northern Region. Its water source is a large reservoir fed by the Bontanga River. MiDA rehabilitated an area of 495 hectares, of a maximum potential area of 800 hectares. The anchor farmer, Solar Harvest, was expected to cultivate an area of 315 hectares adjacent to the irrigation scheme with the intent of pumping irrigation water for its farm from a collection point that is being constructed at the extreme end of the main canal serving the Bontanga small farmers.

Golinga: The Golinga irrigation scheme was originally built in 1965, with a planned capacity of 100 hectares of irrigated land. The scheme draws water from a small reservoir fed by the Jolo River. Currently, the scheme covers of a total area of 65 hectares, with a net area under cultivation of 60 hectares. MiDA rehabilitated an area of 40 hectares. Due to the relatively small size of the Golinga scheme, there is no large commercial anchor farm located nearby. However, the anchor farmer at the Bontanga irrigation scheme was expected to negotiate supply contracts with the Golinga farmers as well.

Construction of the two northern irrigation schemes began on March 15, 2011. Two years later, farmers were getting water but not paying the full cost of water delivery, operation, repair and maintenance. Since no private operator was secured, GIDA was managing the system.

Evaluation Hypotheses and Impact Indicators

Due to poor scheme management combined with inadequate maintenance and repair, the performance of the Bontanga and the Golinga schemes had progressively declined and at the time of the Compact signing, both schemes were operating at less than half their design capacities. These problems had worsened by the refusal of the farmers to pay the full amount of the assessed irrigation fees. In recent years only about 300 hectares had been cropped in the dry season, with a similar crop area cultivated during the rainy season. During the dry season, the poor condition of the irrigation canals would limit the amount of water that can be provided for crop production. During the rainy season, poor soil drainage caused by clogged and silted drainage canals made a large part of the scheme area too wet to cultivate.

The work being carried out by MiDA sought to rehabilitate the existing network of both irrigation and drainage canals, which would increase their operating efficiency and improve the overall performance of the entire scheme. The rehabilitation would also equip the scheme with monitoring equipment for better control and efficient use of irrigation water. The work carried out by MiDA was expected to make it possible for the small farmers on the schemes to cultivate their crops without regard to rainfall patterns, which would substantially increase their production output. Furthermore, contract farming arrangements with the anchor farmer would provide a reliable market outlet for the small producers, as well as access good-quality seed and to crop production technology.

The main expected hypotheses regarding outcomes of the irrigation activity were:

1. With irrigation water supply and good crop management, it was to be possible to grow three crops per year in both the north and south regions. Previously, with the limited availability of irrigation at Botanga and Golinga during the dry season, and given the inadequate farm drainage that limits the use of land in lower producing areas during the rainy season, one to

two crops annually was the norm for farmers in both locations. More crops would translate into increased farm production/output and greater farm incomes for small farmers.

2. Irrigation makes it possible to cultivate a mix of higher-value crops. For example, high-value vegetable crops, particularly leafy vegetables are greatly susceptible to losses from pests and diseases during the rainy season. With irrigation, these crops could be successfully grown during the dry season when prices tend to be higher. A higher-value crop mix provides greater farm incomes for small farmers.

3. Irrigated, commercial crop production is much more intensive and has a greater labor requirement than subsistence agriculture. Labor requirements for the anchor farms will be substantial – up to seven workers per hectare. In addition even small farmers could need to engage hired labor, especially for harvesting crops. The resulting employment generation at the irrigation schemes could be substantial.

Based on these hypotheses we propose to use the following indicators to measure impact:

1. Total annual household income
2. Total annual household income from crop production
3. Paid employment per household
4. Crop mix: Annual production output (kilograms) for each of the five most important crops produced per household. We aim to measure the changes form low to high value crops.
5. Crop yield: A crop will be selected as a representative at each irrigation site and its output per unit area (kilograms per hectare) will be monitored for each crop cycle.

Evaluation Methodology

We propose a double difference estimate with matching and covariates controls as appropriate to evaluate the impact of irrigation activities on small farmers. The treatment group for each irrigation scheme is comprised of small farmers that belong to FBOs that operate within the geographic perimeters of the irrigation scheme, who receive irrigation. The control group is composed of the farmers outside the water supply perimeters who do not receive the benefits of the irrigation schemes but are similar in characteristics to the treatment groups.

In order to estimate the effect on paid employment we can use both a difference in difference approach and an IV approach, based on a distance indicator, similar to post-harvest activities. If we assume that small farmers who live closer to the anchor farmers are more likely to benefit from an increased demand for labor on anchor farms, we can instrument treatment by using "farmer's distance to the anchor farmer" either as a continuous variable or as discrete categories defined by distance of, for example, 20km radius and 30km radius as suggested before.

Risk Factors

The number of treated farmers in the Golinga and Botanga scheme is only around 150. NORC collected data from this group and created a control group of similar size. We expect this number to be large enough to detect the impact of the activity.

Data Sources

For the evaluation of the irrigation schemes in the Northern area, Bontanga and Golinga, we use data from the MiDA GIS database to geo-locate FBOs and construct borders that separate FBOs/farmers that do and do not benefit from the irrigation water supply. The main source of data for this evaluation is a large-scale farmer survey that NORC conducted in late 2012. In this survey, which covered 656 farmers, we collected data on farmers based on the FBO information lists provided by MIDA. Data from the NORC 2012 farmer survey would provide baseline information for farmers associated with the treatment and comparison FBOs.

A second wave of the NORC farmer survey was to be conducted to construct an endline. The endline survey should ideally take place as late as possible, given that there have been delays in the completion of irrigation schemes, and sufficient lag time (two years or two crop cycles after the irrigation schemes are functional) is required to allow the interventions to show results. The exact timing of the endline data collection is yet to be determined.

Key Considerations

When the compact ended in February 2012, construction of the Kpong irrigation scheme had not been completed. By contrast, the Golinga and Botanga schemes were completed and operational by the end of the compact. They were transferred to GIDA; although farmers are receiving water, they are not paying the full cost of water delivery, operation, repair and maintenance. As we anticipated in our previous design report, there has been insufficient time before the compact ended to organize the scheme management structure and to institute the management and operating systems needed to successfully manage the scheme. No private operator was secured for the two completed schemes, therefore GIDA is managing the system as it always has done without covering full costs. These difficulties could diminish the impact of MiDA's investment in the near term.

Finally, this activity cannot be interpreted as just providing irrigation. The anchor farmer presence is an important and non-random part of the intervention.

B.3. Rural Development Project: Community Services Activity, Education

Background Information

The Education Sub-Activity under the Rural Development Project funded the construction and rehabilitation of schools. Investing in educational facilities is expected to increase student enrollment and attendance and reduce drop-out rates by improving access (reducing travel time) and creating a better learning environment in the schools. Improved access to schools and conditions may also reduce absenteeism among teachers.

The Education Sub-Activity was rolled out in two phases. While waiting for the completion of Needs Assessment and the Environmental Impact Assessment Study Reports, the Community

Services Project Department was allowed to select a few school blocks which were listed in the beneficiary Districts' Medium Term Development Plans and requiring urgent attention for rehabilitation under a limited budget. Phase I of the Education Sub-Activity (2007-2009) was viewed as a quick start project in pilot districts.

Phase II of the Project included construction and rehabilitation of educational facilities in 151 communities. These facilities included the construction of 106 two-unit classroom blocks for the kindergarten level, 41 three-unit classroom blocks for the junior high and primary levels, and 29 six-unit classroom blocks for the primary level. All schools received urinal facilities and a full complement of school furniture (wall-mounted black-board, furniture, desks and benches, teachers table and chairs, classroom cupboards and basic furniture) for all constructed classroom blocks and head teachers' offices. Electrical wiring of the school building was done where electricity is available or where the District Assembly assured the availability of electricity in the short term. Phase II activities started in 2009 and were completed by end of Compact.

Phase I school selection did not follow a systematic selection process or defined criteria. Conversely, however, selection of intervention schools for Phase II followed strict selection procedures, which began with a close review of District Medium Term Development Plans to identify the District Assemblies' list of schools prioritized for construction and rehabilitation. These schools were then ranked and scored according to a preset set of criteria.

School Ranking and Selection Process:

At the National Level: The FBO concentration in each Zone was evaluated, taking into consideration the proportion of the total number of FBO in the Zone and the total number of FBO's under the MiDA project. The Zonal FBO ratio was then derived, and MiDA's budgets, including the Community Services Project budget, were allocated accordingly. Based on this allocation criterion, the Northern Zone received 30 percent of the MiDA education sub-activity funding, the Afram Basin received 50 percent, and the South region received 20 percent.

At the Zonal Level: Two parameters were used to rank Districts. In order of priority they were:

- Poverty index (40% weight)
- Number of FBOs in the District (60%)

At the Community Level:

- The Communities with the highest number of FBO's were ranked in descending order
- Schools with sub-standard structures, Category A
- Schools with inadequate/deficient educational facilities, Category B.

Sub-standard School Structures **--** Priority is given in the following order:

- Classes under Trees
- Classrooms in unsafe structures (Mud, Open Sheds, etc.)
- Uncompleted School Structures
- Schools in rented accommodation
- Schools in unclad Pavilions

Schools with Inadequate/ Deficient Facilities **--** Priority is given in the following order.

- Schools with shortfalls in classroom accommodation
- Schools with Shift System
- Schools without prescribed ancillary facilities, shall be provided with the following in order of priority:
    - Toilet and Urinal,
    - Potable Water Facility
    - Staff Accommodation
    - Library
    - Computer Laboratory
    - Dining Hall
    - Sickbay
- Schools lacking Recreational and Sports Facilities

Schools in each target district were ranked and scored from 0 to 100 according to the community-level criteria, with high priority schools in need of urgent attention receiving a higher score. The final decision on which schools were selected for construction/rehabilitation in a given district was based on this eligibility index and the availability of MiDA funds for the district. Because of funding constraints, not all priority schools in a district were not built/rehabilitated.

Because the Phase II school construction/rehabilitation activities followed a systematic approach that is more conducive to a quantitative evaluation, we focus our impact evaluation only on this phase of the Education Sub-Activity.

Evaluation Hypotheses and Impact Indicators

The hypothesis behind this intervention is that more and better educational facilities can improve educational outcomes. This is can seem obvious for the cases of new construction and outcomes such as enrollment. For example, a new kindergarten facility where there was none would increase enrollment provided that there is demand from parents to send their young children to pre-school. Rehabilitation of existent facilities may have a similar, although more moderate, effects on enrollment.

We also hypothesize that better educational infrastructure creates an environment that is more conducive to learning and staying in school, thereby reducing drop-out rates and increasing attendance.

Access to toilet facilities at the schools can be a benefit for all children and produce positive externalities for others. However, it has been argued that the availability of separate toilet facilities has a larger positive effect on girls' school attendance and enrollment, although empirical evidence is not conclusive on this point.

Based on the aforementioned hypotheses and available data, we propose to use the following indicators for measuring impact:

- Gross enrollment rate (GER) in the catchment areas, total for all schools. Total, by gender, by age groups
- Net enrollment rate (NER) in the catchment areas, total for all schools. Total, by gender, by age groups

- School gross attendance rates: grade-wise number of children enrolled. Total, by gender, by age
- Average % of days that students attend school. Total and by grade and gender.

It is important to note here that we propose to measure enrollment rates for the entire catchment area of the school, and not for the intervention or comparison school in the sample. We do this because the presence of a new school could well draw students away from other schools in the same area, creating a situation in which the new or improved school's enrollment increases, while older, less attractive school facilities lose students and suffer enrollment losses. In such cases the enrollment rate of the school catchment area may remain unchanged. Using catchment area enrollment rates will allow us capture this dynamic process and avoid overestimation of the impact.

Evaluation Methodology Proposed in the Design Report

As described above, MiDA used an eligibility index to select the school units that were reconstructed or rehabilitated in a given district. Schools that received high scores (i.e. schools in dire need of attention) were selected for treatment with available funding; schools below the district-specific cut-off did not receive treatment because the available budget did not permit it. This selection process lends itself to using a Regression Discontinuity Design (RDD) to evaluate the impact of the Education Sub-Activity. This strategy exploits the discontinuity around the cutoff in the eligibility index to estimate counterfactuals. The assumption is that eligible schools with scores just above the cutoff are very similar to eligible schools with a score just below the cutoff that were not selected because of funding constraints. Therefore schools that were not selected for the program but were close enough to the threshold can be used as a comparison group to estimate the counterfactual.

As a first step, we proposed an analysis of the baseline data to test the validity of the RDD, comparing indicators for eligible schools with scores just above the cutoff to eligible schools with a score just below the cutoff that were not selected because of funding constraints. In the RDD literature this analysis is done graphically and allows us to assess how the similarities, or lack thereof, in the two groups of schools. We also proposed to check the school data to determine whether we face a case of sharp discontinuity (SRD) or a fuzzy discontinuity (FRD) and plan our regressions accordingly. In the FRD design, the probability of receiving the treatment does not need to change from 0 to 1 at the threshold. Instead, the design allows for a smaller jump in the probability of assignment to the treatment at the threshold[3].

Data Sources

School level data is collected each year through the Education Management Information System (EMIS). NORC has in its possession the EMIS database corresponding to 2008/09, which can be used as baseline data. In the future, we would need at least a new wave of this survey data to analyze post intervention indicators and evaluate the impact of the education activity.

Other data sources that would be required for this evaluation are the following:

[3] See for example, Imbens, G.W., Lemieux, T., Regression discontinuity designs: A guide to practice, Journal of Econometrics (2007)

- 2010 Census data, which is still in the process of being compiled and, hence, unavailable to us at this juncture. NORC submitted a request for this data to the Ghana Statistical Services through formal channels, and tried to also enlist MiDA assistance in obtaining the dataset when it is finalized. We propose to use age and sex specific population numbers from the Census data to calculate enrollment rates in school catchment areas for the impact evaluation of the Education Activity.

- Key pieces of data that will need to be collected at the district-level Education Offices. This information includes communities (or enumeration areas) that fall into the catchment areas of MiDA and comparison schools; other schools in each catchment area; and other education projects that have been or are being conducted in the districts. The District Education Offices may also be able to provide school-age population figures for each school catchment area; this may be a more direct way to gather this data, rather than estimating it based on Census figures. Collecting education information from 30 districts is a daunting task and relying on overburdened district education officers to fulfill this data collection task is not a recipe for success. Therefore, we proposed exploring the option of hiring a local consultant for a 2-3 month period to visit district offices and schools and collect the requisite data.

Risks

Our School RDD sample consists of 302 schools. We selected this sample using the sample frame provided to us by Lambda Consulting, the consulting group that ranked and selected schools for MiDA interventions. Schools in each district were ranked according to a pre-determined set of criteria and the "most eligible" were selected based on each district's budget for the school intervention. NORC selected treatment and control schools, using the cut-off point for each district – schools with ranking scores immediately above the cut-off constitute the treatment group and those immediately below the cut-off constitute the control group in the RDD.

In order to evaluate the Education Activity, per the RDD methodology described above, we need to:

1) Match our sample schools to at least two waves of EMIS data (since data for several outcome indicators will be drawn from the EMIS dataset)
2) Match schools zones/localities to zones/localities in the census (as described above, to calculate enrollment rates in the school zone or area; for the evaluation we wish to measure the impact of the intervention on enrollment in the entire school area or zone, and not just in the school that was rehabilitated. This will allow us to capture and account for any movement of students away from non-rehabilitated to improved schools)

For the matching of sample schools to schools in the EMIS datasets, we have found that we cannot use school names or addresses as we were planning to do, because the school information we have from MIDA is largely incomplete and does not line up with school names in the EMIS dataset. Therefore, we will need to use EMIS codes to do this matching. Unfortunately, the school lists provided to us by Lambda Consulting do not contain the EMIS codes, and we have been unsuccessful in our multiple attempts to reach Lambda Consulting staff. Therefore, we would have to obtain EMIS codes for our sample schools directly from the schools, either by calling the school or, in the majority of cases, visiting individual schools. Sample attrition associated with this matching process could be significant and, below, we present two such scenarios, assuming different rates of attrition:

Scenario 1 – Successful matching process

- Locate 100% of the schools in our sample;
- Obtain 95% of the EMIS codes; and
- Match 95% of the schools to each wave of EMIS data
- AND
- Match 95% of school localities to the census localities,

The final sample will consist of 245 schools

Scenario 2 – Less successful matching

- Locate 95% of the schools;
- Obtain 90% of the EMIS codes; and
- Match 90% of the schools to each wave of EMIS data

AND

- Match 90% of the schools localities to the census localities,

The final sample will consist of 188 schools

Even a successful matching of the data most likely will produce a small final sample. We are concerned that this sample will not be large enough and the quantitative analysis will lack precision and will not produce statistically significant results.

In addition we are concerned about the availability of the census data at the level of detail that would be needed for the RDD. We know the detailed information at the appropriate geographic level exists for Census 2000 (which was released in 2005-2006), but it is not clear whether that same level of detail will be available for in the 2010 census. Although we had hoped to have access to the 2010 census data by now, it has not been released yet.

Alternative: Qualitative Approach

This approach will focus on collecting qualitative information from a few communities where school construction took place and some communities that applied for the investment and were not selected. The goal is to conduct a qualitative evaluation using key informant interviews and focus groups with parents, teachers, school and district authorities and students. We anticipate conducting FGDs and KKIs in approximately 12 communities in 6 districts, at a rate of 3 FGDs and 4-5 key informant interviews per community. We propose to conduct 60 minute (on average) FGDs with teachers, parents, and students, with each group consisting of eight participants. KIIs will target school principals and administrators.

There is no question that new and better facilities are desirable in general, but through the qualitative research we would try to find out if those new or better facilities have translated into increased enrollment, less absenteeism, if availability of toilets changes girls' attendance, etc. If the investments did not translate into better outcomes, we will try to identify the reasons.

We would similarly gather qualitative information about trends in enrollment, absenteeism, attendance among girls, etc. in non-intervention communities where the schools are –measured by the ranking used by MIDA to allocate interventions- very similar to the selected ones. This would constitute a form of counterfactual that will allow us to have a better understanding of the situation.

This approach does not constitute a formal IE but it can be quite informative and could help to design future activities in the education area.

B.4 Rural Development Project: Financial Services

Background Information

MiDA's Financial Services Activity sought to establish a computerized networking system (through WAN using VSAT) between rural banks and the Apex Bank Server (financed by the project). The activity was aimed at improving financial service delivery, operations, and access to information at rural banks with the objective of enhancing the depth and value of rural financial services and widening access to savings services and cash transfers.

Computerization and Connectivity: To date, all Rural Banks (134) in Ghana have received computers and VSAT satellite dishes, and have connected to the server for fully automated operations. In addition to the computers and satellite dishes banks received a full commercial banking software package (eMerge) that enables them to access computerized front and back office applications for real time transactions with their customers, track cash flow, revenues, and expenses by profit/cost center, and update customer accounts with an easy end of the day processing. By the end of the MCC Compact, all 134 rural banks had received the full package of upgrades. Furthermore, all rural banks had received a scanner and software for the Check Codeline Clearing System.

According to information received from the Project Management Team, the Financial Services Activity used very strict prioritization criteria for the migration of the rural banks onto the eMerge platform. The guidelines for the prioritization were designed by the project management team (PMSC) and were deliberated and amended where necessary by the project technical committee for final approval by the project steering committee. The criteria however got changed several times as necessitated by the exigencies of the situations including the tight project schedule, limitation of the data center infrastructure, and the location of RCBs. The key prioritization criteria were as follows: availability of correct and balanced data (weighted highest), infrastructure readiness, information security compliance, and basic computer appreciation and training for staff. Unfortunately no registry of the selection criteria and process was kept.

Once fully automated, all financial data will be stored centrally at the APEX Bank. Rural Banks use the software by connecting to the Apex Bank's Server. The accounting and banking data is therefore available in real time at the Rural Bank and at the Apex Bank. The Rural Banks are thereby enabled to provide accurate, up-to-date, real-time statements to their customers.

Cheque Clearing and Cash Transfers: All rural banks have received a scanner and software for the cheque code clearing (CCC). However, due to the absence of data on cheque clearing times, the evaluation can provide only a qualitative assessment and not a quantitative measurement of the impact of this activity. Nevertheless, one of the key indicators (see below) to be analyzed quantitatively would be "non-interest income," which includes income from cheque clearing (one percent of the amount of the cheque).

The project did not finance a specific facility for cash transfers. The Apex bank has developed the Apex link software through internal resources that will be built on the platform of VSAT and computers financed by the project. The cash transfer module is expected to become live in 2012. Other cash transfer systems available at Rural Banks include the *e-zwitch*, a biometrically recognized debit card that can be used for loading and unloading funds to the same card or from one card to another card. The Rural banks have a POS with a SIM card that transfers the information using the wireless telephone coverage. Irregular coverage and poor service has hampered the effectiveness of this product, which is no way competing with any of the services provided by the MiDA Financial Services Activity.

Evaluation Hypotheses and Impact Indicators

This intervention intended to improve the efficiency of financial transactions at rural banks and make it more attractive for people to use the banking system. The computerization of operations and connectivity that MIDA has supported for the rural banks should have the following effects:

- Improvement of the speed and reliability of transactions
- Improvement accuracy and availability of accounts information
- Reduction in check clearing times

These improvements not only make bank operations more efficient and potentially less costly per transaction, but they also improve the customer experience with the services offered by the bank as well as reducing the incidence fraud. This, in turn, is expected to increase the number of clients, the number of deposits and credits, and the number of bank operations, among others.

The following indicators cover the key areas of operating efficiency, transaction costs, number and volumes of various types of deposits and changes in the customer base as compared to non-computerized rural banks.

i. Number of all accounts combined (savings deposits, demand deposits, time deposits, Susu)
ii. Number of all deposit accounts (savings deposits, demand deposits, time deposits)
iii. Number of savings and fixed deposit accounts
iv. Value of all accounts (savings deposits, demand deposits, time deposits, Susu)
v. Value of all deposit accounts (savings deposits, demand deposits, time deposits)
vi. Value of savings and fixed deposits
vii. Calculate average deposit size: v/ii
viii. Calculate average size of savings and fixed deposits: vi/iii
ix. Calculate ratio of fixed and savings deposits to total deposits: vi/v
x. Financial expenses (interest payments on deposits and other borrowings by RCB)
xi. Operating expenses (excluding financial expenses)
xii. Total expenses: x + xi

xiii. Interest income
xiv. Non-interest income (commissions and fees plus other income)
xv. Net financial income (interest income minus interest expenses): xiii-x
xvi. Total income from all sources (non-interest and interest income): xiii+xiv
xvii. Net Income: xvi-xii
xviii. Calculate ratio for non-interest income to total income: xiv/xvi
xix. Calculate OER or Operating Expense Ratio: xi/xvi
xx. Calculate operating costs per customer: xi/i
xxi. Staff expenses
xxii. Calculate ratio of staff expenses in total expenses: xxi/xii
xxiii. Calculate ratio of operating expenses to net financial income : xi/xv
xxiv. Banks's total adjusted capital from all sources including reserves
xxv. Calculate profitability ratio (ROE): Net Income divided by RCB's adjusted capital from all sources: xvii/xxiv
xxvi. Total adjusted Assets in the B/S
xxvii. Calculate return on adjusted assets (ROA): xvi/xxvi

Due to the lack of data on check clearing times, we will not be able to estimate quantitatively the effect of the activity on the speed of clearing. It could be done qualitatively by interviewing bank managers.

Evaluation Methodology

We propose to use a difference-in-difference estimator (with matching or covariates controls when appropriate) where the staggered rollout of the intervention (VSAT, computerization and connection to the server) can be used to define treatment and comparison groups. Because rollout occurred in batches, the treatment group would be comprised of banks that received the treatment early and the control group would include the rural banks that receive the treatment later.

The first bank was connected to the server in June 24th, 2010 and in the following months other banks followed. By July 25th, 2011 a total of 60 banks were connected to the server. We will consider this first group of connected banks our "early treatment" group. It was not until October 17th, 2011 that the first bank in the second batch received a connection and the process of linking the rest of the banks to the server is still going on. The second batch of banks could make up the control group.

Originally, we had proposed to do two separate analyses, one estimating the effect of distributing the equipment to the banks and a second one to estimate the additional effect of connecting the rural banks. Unfortunately, the distribution of equipment is not sufficiently spread out over time to allow creating treatment and comparison groups. In the case of connection to the server, a time lag exists that makes the approach feasible, as half the banks were still not connected a year after the first banks went live.

Although using two groups of banks, early treatment and later treatment or control, is appealing because of its simplicity, it is not the only way to approach this evaluation. Using the monthly data for each individual banks, the general difference-in-difference models can be specified as a fixed effects regression model for panels:

$y_{it} = \alpha D_{it} + \beta \mathbf{x}_{it} + \lambda_t + \mu_i + \varepsilon_{it},$

where y_{it} is the outcome of interest, $\mathbf{x}_{it}$ is the vector of the subset of control variables in the vector $\mathbf{x}$ that vary both across units and time, μ_i is a time-invariant effect unique to unit i, λ_t is a time effect common to all units in period t (month and year), and ε_{it} is a unit time-varying error distributed independently across units and time and independently of all μ_i and λ_t and finally D_{it} indicates treatment.

Alternatively we can perform this analysis in annual differences in order to remove seasonality such that

$\Delta y_{i\,y} = y_{it} - y_{it-12}$ where y denotes year

and estimate the effect of the variable "time since connected to server" on changes in the proposed indicators, while we control for covariates.

Risk Factors

There are two main complications that can threaten the feasibility of the proposed methodology. The first is the number of observations. There are only 134 rural banks and all of them were eventually treated. With such a small sample, it is possible that no effect may be detected when using the two group difference-in-difference approach. We therefore propose to complement the analysis with the fixed effects approach described above.

The second difficulty we face is that rural banks were selected for migration to the eMerge platform according to a set of prioritization criteria (availability of correct and balanced data, infrastructure readiness, information security compliance, basic computer appreciation for staff, etc.). The original priority order is not available to us. Although we learned that the order was altered because of different circumstances, we are concerned that the rural banks that went live earlier are also 'superior' in other unobservable dimensions that could introduce bias in our estimates. If this is the case, it is likely that the estimation of the activity will be biased upwards.

Data Sources

Having reviewed existing data sources, we became aware that the data necessary to estimate indicator of interest should be available at the ARB APEX Bank in the form of monthly returns from individual rural banks (not by branch). The monthly returns for each bank may be available in paper format for the period prior to full automation of a given bank, and electronically for the period following full automation of the bank.

In 2011 NORC initiated a request to access these data. Initial conversations with APEX staff were encouraging. MIDA sent a formal request for the data and obtained the agreement from APEX to collaborate in NORC's efforts to obtain data necessary for the evaluation. However, as of May 25, 2013 we had yet to receive the electronic (post-automation) data, despite repeated attempt to obtain it directly from APEX. The APEX manager authorized to provide NORC with the electronic data has been on extended leave and all requests have been unanswered. Currently MCC is trying to gain access to this electronic data, without which the proposed evaluation will not be possible.

The process of obtaining paper records of monthly bank returns for 134 banks for a period of several months is a non-trivial undertaking. These paper returns, which are currently stored in archives in Tamale, will have to be scanned or photo-copied, and the necessary information from these paper reports will then have to be entered into an electronic database.

B.5 Agriculture Project: Credit Activity

MCC is currently deciding on whether or not an evaluation of MiDA's credit activity is to be undertaken by NORC, taking into account the fact that several assessments and an independent audit of this activity have already been undertaken.

C. DATA SOURCES

In this section, we discuss the data sources that we have identified as useful for the evaluation. Since our initial proposal, we have had the opportunity to review various datasets mentioned in the RFP, as well as identify new data sources. Below, we present each possible data source and discuss the pros and cons of each.

C.1 GLSS5+ Database

The GLSS5+ was a large-scale household survey conducted in 2008 that was funded by MiDA and covered many of the MiDA target districts. Specifically, 9,300 households in 27 Enumeration Areas in the original 23 program districts were surveyed. As such, we anticipated using this data source heavily for the impact evaluation of several of the MiDA activities and sub-activities.

In November 2011, NORC obtained the dataset and supporting documentation, and met with researchers from ISSER who were in charge of the survey. In the course of reviewing the data and discussing its contents with ISSER staff, we have become aware of two significant limitations to using the data for the evaluation.

1. GLSS5+ was designed to be used at district level, but not at more disaggregated levels such as communities or Census enumeration areas. The activities and sub-activities that NORC is evaluating, however, will have impact at lower levels, such as communities, FBO coverage areas, and school catchment areas. It is unlikely that the impact of these activities, particularly given the sample sizes available, would be detected at the district level.

2. The GLSS5+ household data does not include more specific information about the location of the household than cluster. As such, it is not possible to link households with the communities or schools catchment areas in which they are located. By extension, it is also not possible to link GLSS5+ households with MiDA intervention sites and boundary areas obtained from the MiDA GIS dataset. It is essential to be able to create correspondence between individual farmer data and MiDA GIS data in order to identify treatment households and potential controls. GLSS5+ data only allows matching at district level.

Given these limitations, we do not anticipate using the GLSS5+ data for this evaluation..

C.2 FBO Survey Data

The FBO survey, funded by MiDA, was fielded by ISSER during 2008. It covered 2,928 farmers in its first wave. The FBO collects data among farmers on demographic characteristics, education and skills/training, health, activity and occupation, migration, housing and housing conditions, assets, land ownership and land transactions, agriculture production, and non-farm household enterprises.

Currently, we have two versions of the FBO database: one provided by MCC, and a second one was provided by ISSER. Neither version is complete. For example, the ISSER's version is missing some sections of the survey for Phase II FBOs; and the MCC version is missing data on variables such as phase of FBO training, FBO batch number, and degree of technological adaption. We are currently working on obtaining a complete version of the FBO survey from ISSER.

In theory, the FBO Survey dataset could have been used as a baseline for the impact evaluation, and our preceding discussion of the evaluation designs for the agriculture activities assumes that we will do so. However, at this stage, the dataset is missing some key variables that are critical for the evaluation.

1. The only way to match the FBO data with the MiDA intervention sites in order to select a sample of treatment and control farmers for the evaluation is by using the geo-coordinates for the FBOs that are available in the MiDA GIS dataset. Although both datasets contain FBO IDs, these identification codes are different in the two datasets and, therefore, cannot be linked. FBO names are also quite different in the two datasets due to different versions of the names, different spelling, and abbreviations. We have found that we can do a fairly good job of matching FBOs with MiDA intervention sites by using geographical coordinates. However, both of the versions of the FBO survey we have include only the coordinates for Phase I FBOs. It is essential that we obtain the coordinates for Phase II FBOs as well if we are to use the FBO data in our evaluation. We have contacted ISSER about obtaining this missing data and will request MiDA's help in facilitating this data acquisition.

2. Some sections in the dataset corresponding to Phase I (in both versions) do not contain variables specified in the survey instrument. In particular, questions contained in Section 10, Part G - Output Details, Main Season and Minor Season of the questionnaires do not have corresponding variables in the dataset – for example, quantity of crop produced, revenues, and post-harvest losses, among others. This section in the survey captures crop harvest during the main and minor seasons and, as such, is important for the evaluation of irrigation and post-harvest activities.

3. Reports produced by ISSER contain information about crop yields. It is not clear from the survey instrument which questions were used to calculate this indicator, as there seems to be a mismatch between data and instrument. Crop yield is a very relevant indicator for the evaluation of the agricultural activities, and we would like to be able to compute it and use in our impact evaluations. We will work with ISSER on this issue, but at the present it is unclear if such an indicator is available at baseline or not.

Because of the limitations of this data, NORC undertook its own baseline data collection among a sample of treatment and control farmers linked to the Golinga, Botanga, and Kpong irrigation schemes.

C.3 Baseline NORC Dataset

Given the limitations of the GLSS5+ and FBO datasets, MCC and NORC decided that a primary data collection effort for the ABC and irrigation evaluation was warranted.

NORC collected data from farmers around irrigation schemes and ABCs. Some of those farmers are supposed to be directly affected by the activities (treatment groups) while others are similar in characteristics but farther away from the interventions and less affected or not affected at all (control groups). Data was collected on socio-demographic characteristics, household composition, participation in FBOs, agricultural activities, plots, crops, irrigation, production, relationship with ABCs, assets, hunger, and employment, among other indicators.

C.4 MiDA GIS Database

We have obtained the GIS Database from MiDA, which contains shape files and data for MiDA interventions. While this dataset has proven to be very useful in geo-locating intervention sites and matching them with FBOs, we have discovered two problems to date. The first is that FBOs IDs do not match those of the FBO survey. However, with some effort, we were able to match the FBOs and farmers in the two datasets by using the geographical coordinates found in the FBO survey data.

The second problem is that there are several missing data points. For example, shapefiles containing ABC locations were not included with the rest of the GIS data we received from MiDA. We worked with MiDA to resolve this issue, and were able to obtain the position of 8 ABCs. Although we tried numerous times MIDA did not supply the location of the other 2 ABCs.

C.5 APEX Data

We have identified several indicators that are suitable for analyzing the impact of computerization and connectivity of rural banks. The indicators, listed in Section D.2, cover changes in banks' operating efficiency, transaction costs, number and volumes of various types of deposits, and customer base. This data, as mentioned before, may be available in banks' monthly returns in either paper or electronic form.

Specifically we would need:

1. From the Efficiency Monitoring Unit at APEX, as discussed with Mr. S. Twumasi Ankrah, copies of the complete monthly paper returns for the 134 rural banks starting from January 2009 through month of automation (which would vary by bank).

   We would like to obtain copies of these paper records as soon as possible. However, it makes no sense to collect this data and incur costs if the electronic data is not collected as well (see below). The paper returns are currently stored in the city of Tamale. We will need to capture these data in electronic format.

2. From the Data Centre at APEX, as discussed with Mr. Michael Appiah, an excel dataset that includes all reporting variables linked to the indicators listed above for the period following automation for each of the 134 banks. This list of variables includes the following:

   - # of customers using demand deposits
   - # of customers using savings deposits
   - # of customers using time deposits (fixed deposits)
   - # of customers using Susu deposits
   - # of customers getting loans
   - # of customers with overdrafts
   - total # customers
   - # of fixed deposits
   - # of demand deposits
   - # of savings deposits
   - total value of demand deposits
   - total value of savings deposits
   - total value of fixed deposits
   - total operating cost (expense)
   - non-interest income (commissions and fees)
   - total income

Without these data, the impact evaluation we propose for the financial services activity will not be possible.

At present, we are still working with MCC to obtain the data. We note that an agreement had been set in place an agreement with APEX to ensure that NORC has access to this data, however the transfer of data did not take place yet.

C.6 School Census, EMIS Database

School censuses are annual and information is collected through the Education Management System (EMIS). We have identified this database as the main source of data for school-level indicators to be use in our impact evaluation. For this purpose, we will need to have a minimum of two waves of the data, one to be used as baseline and another as endline.

NORC obtained the EMIS for the 2008-2009 school year through our own contacts, and this will serve as a good baseline for the evaluation. It is important to note, however, that we were unable to access this data through the Ministry of Education. It is critical that we make arrangements through MiDA and MCC to ensure that we will be able to obtain the EMIS dataset for the endline, for the years between 2010 and 2014.

C.7 Population Data from Ghana Statistical Services

Through the EMIS 2010 database, we have access to school-level enrollment numbers. In order to calculate area-specific enrollment rates, we need age and sex specific population data for enumeration areas that fall within each school's catchment area. We hope to obtain this data from the 2010 Census when it is available. We have already made a formal request to Ghana

Statistical Services, however we have not received a response after almost 2 years. We would like to request MCC's help in accessing this data as soon as the final dataset is available.

In order to compute enrollment indicators for the endline we will need population estimates for 2011-2012 or 2012-2013 based on the census 2010, if they are available. This data is usually computed by the statistical institutes, GSS in our case. Otherwise, we will need to use 2010 census data as the base and estimate endline indicators under the assumption that population growth rates are the same across geographical areas of interest.

C.8 School District Data

We have identified key pieces of information that we would collect at district-level Education Offices, namely, communities that fall into the catchment areas of MiDA and comparison schools, other schools in each catchment area, and other education projects that have been or are being conducted in the districts. The District Education Offices may also be able to provide school-age population figures for each school catchment area; this may be a more direct way to gather this data, rather than estimates based on Census figures. We propose to hire a local consultant to collect necessary data at the district level.

ANNEX 1: PROPOSED EVALUATION METHODOLOGIES

1. Difference-in-Difference

In this approach we compare changes in outcomes over time between a population that has been treated (treatment group) and a population that has not (comparison or control group). The first difference (the before-and-after) for the treatment group controls for factors that are constant over time in the group. The second difference (treatment vs. controls) aims to control for time varying factors.

The general difference-in-differences estimator is given by:

$$\hat{\varphi}=E(y_{t1} - y_{t0} \mid \mathbf{x}, D=1) - E(y_{c1} - y_{c0} \mid \mathbf{x}, D=0)$$

where,
y = outcome variable
x = observable independent variables
t = treatment group
c = control group
0 = baseline or beginning of study
1 = end of study

Note that treatment and comparison groups do not need to be identical at baseline, however changes in the comparison group should represent what would have happen to the treatment group in absence of the intervention. Difference-in-difference can take care of differences in the treatment and control groups (observable and unobservable) that are time invariant, but does not address the differences that change over time.

The general difference-in-difference models can be reduced to the fixed effect model if the expected conditional y variable only differs by a constant α

$$y_{it} = \alpha D_{it} + \beta \mathbf{x}_{it} + \lambda_t + \mu_i + \varepsilon_{it}$$

where $\mathbf{x}_{it}$ is the vector of the subset of control variables in the vector **x** that vary both across units and time, μ_i is a time-invariant effect unique to unit i, λ_t is a time effect common to all units in period t, and ε_{it} is a unit time-varying error distributed independently across units and time and independently of all μ_i and λ_t (see Galiani, 2002, Chamberlain, 1984; and Heckman and Robb, 1985).

Difference-in-difference can be combined with matching. When using propensity score matching (several different approaches to PSM exist), we try to find units in the control group that are the closest to the treated units by computing the probability that a unit will belong to the treatment based on the observable characteristics. The goal is to mimic a randomized assignment when it does not exist. External validity requires that we find matches for all treatment units, i.e. that we find common support.

We plan to use this approach to evaluate the impact of several activities such as ABCs, Public Pack Houses, Cold Rooms, Irrigation, and Financial Services.

2. Instrumental Variables

In addition to the difference-in-difference approach, we have also proposed using Instrumental Variables. A naive OLS estimator would be biased if the explanatory variables are correlated with the error terms. This is likely to be the case in most of the activities we want to evaluate given that participating FBOs where selected into treatment based on some characteristics unobservable to us. However, if an instrument is available we may obtain consistent estimates. Our instrument must be correlated with the endogenous explanatory variable (treatment, in our case) and uncorrelated with the error term.

We propose to use "distance to facility" as an instrument for treatment given that in our case the location of the farmer is not endogenous. The rationality behind this idea is that some farmers are more likely to participated into treatment given that they are located closer to a particular facility, for example an ABC.

We calculate IV estimates using two-stage least-squares (2SLS). In the first stage, treatment (T) is regressed on all the exogenous variables in the model, Z. The predicted values for T^* are obtained and in the second stage we regress:

$$Y_i = \beta\, T_i^* + \varepsilon_i$$

3. Regression Discontinuity Approach

In the case of the education activity, there is an eligibility ranking and clear cut-offs by district to assign schools to treatment and control. We have, therefore, proposed a Regression Discontinuity Approach. The RDD strategy exploits the discontinuity around the cut-off score to estimate the counterfactual. The idea is that units close to the threshold, just above and just below the cut-off, are similar.

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2(X_i - X_c) + \varepsilon_i$$

where:

Y is the outcome variable
T is the treatment dummy
X is the assignment variable
X_c is the cut-off
β_2 predicts outcome from assignment
β_1 is the estimate of treatment effect