Nepal
Agriculture and Food Security Project (AFSP)

Impact Evaluation Concept Note (Draft)

Development Impact Evaluation (DIME)[1]

August 2013

[1] This concept note was prepared by Daniel Stein (DIME) and Ritika D'Souza (DIME). It benefitted from helpful additions and comments from Maria Jones (DIME) and Soham Sen (SASDS)

Table of Contents

Acronyms

AFSP—Agriculture and Food Security Project

BCC – Behavior Change Communication

FG- Farmer Group

FNSE -- Food and Nutrition Status Enhancement (component of AFSP)

FFS – Farmer Field Schools

GAFSP – Global Agriculture and Food Security Program

GoN – Government of Nepal

ICC- Intra-Cluster Correlation

IE – Impact Evaluation

IYCF – Infant and Young Child Feeding

KISAN- Knowledge-based Integrated Sustainable Agriculture and Nutrition (USAID Project)

MDES- Minimum Detectable Effect Size

MoAD – Ministry of Agriculture Development

MoHP – Ministry of Health and Population

NLSS – National Living Standards Survey

PMU – Project Management Unit

TDA—Technology Dissemination and Adoption (component of AFSP)

1. Executive Summary

This Concept Note presents the ongoing impact evaluation (IE) of the GAFSP-funded Agriculture and Food Security Project (AFSP), part of the World Bank's Development Impact Evaluation (DIME) partnership with the Global Agriculture and Food Security Program (GAFSP). The AFSP is an innovative program designed to improve the livelihood of crop and livestock farmers by developing and facilitating the adoption of improved inputs and technologies. It also aims to improve the nutritional practices of mothers and young children.

The IE is a multi-arm randomized controlled trial that estimates the impact of AFSP's investments in technology development and dissemination on farmer livelihoods. It provides an important contribution to understanding the interaction between agriculture and food security, by measuring the relative impact of both supply-side (improved seeds, kitchen gardens, backyard poultry) and demand-side (behavioral change communication) interventions on nutritional outcomes of pregnant women and young children. This will contribute to global knowledge on how to best leverage agriculture projects to ensure improved nutritional outcomes.

1.1 Strategic Context

Over the last 15 years, Nepal has made significant progress in poverty reduction; poverty rates fell from 45 percent in 1995-96 to 25 percent in 2010-11. In spite of this, food and nutrition security remain challenges for the country. Due to low usage of improved seeds, chemical inputs, and irrigation, Nepal's agricultural productivity is among the lowest in South Asia.[2] The Nepal Thematic Report on Food Security and Nutrition 2013 found that one in five households had an inadequate diet based on the Food Consumption Score, and one in four households were considered food poor.[3] Of children under five, 41 percent are stunted, 11 percent are wasted, and 29 percent are underweight. The Mid and Far West regions (where AFSP will be implemented) have chronic food deficits.[4] Such high rates have dramatic economic costs since adequate nutrition contributes to productivity, economic development, and poverty reduction and heath by enhancing physical work capacity, cognitive development, school performance, and reducing disease and mortality.

---

[2] Nepal Thematic Report on Food Security and Nutrition 2013, Nepal Planning Commission Central Bureau of Statistics, http://documents.wfp.org/stellent/groups/public/documents/ena/wfp256518.pdf

[3] Nepal Thematic Report on Food Security and Nutrition 2013, Nepal Planning Commission Central Bureau of Statistics, http://documents.wfp.org/stellent/groups/public/documents/ena/wfp256518.pdf

[4] Ibid

The Government of Nepal (GoN) is making a concerted effort to improve food security and sustain economic growth. In 2010, the GoN led consultations with donors, civil society organizations and other stakeholders to develop a Country Investment Plan to comprehensively address agriculture and food security, including issues of food availability, access and utilization. Building on this, GoN was awarded a competitive grant by the Global Agriculture and Food Security Program (GAFSP) to implement the Agriculture and Food Security Project (AFSP).

1.2 Impact Evaluation Overview

The AFSP IE focuses on the impacts of the technology adoption and nutrition enhancement components of the project. The causal effects of these components will be identified using a randomized phase-in of project components at the VDC level.[5] A sample of VDCs outside of the project area will also be used as additional controls.

The evaluation will test two major features of the AFSP project. First, it will measure the effect of AFSP's agricultural initiatives that on yield, income, and nutritional practices. This includes nutrition-specific interventions such as kitchen gardens and backyard poultry. It will also measure the effect of Behavior Change Communication (BCC), which is designed to increase the demand for nutritious food. By separately phasing in the supply-side and demand-side interventions, the evaluation will be able to measure their individual effects, along with their interactions.

The evaluation will also test program variations in BCC messaging as well as mechanisms to deliver BCC. With regard to program variations, the IE will compare the impacts on behavior change from generic nutrition education messages against those adapted to gender and social contexts. With regard to delivery mechanisms, it will test the additional impact of providing BCC messaging to individual mothers via peer counselors in addition to providing such messages to a group. The effects of these variations will be measured by randomizing their implementation at the VDC level.

1.3 Capacity Building and Research Impact

The DIME team has been working closely with operational and program management staff from both the GoN and the World Bank (WB) from the inception of this IE. The team designed the IE framework during DIME's capacity building workshop in Naivasha, Kenya (April 2012). During this workshop, the AFSP team was trained on IE evidence and methods through case studies on the use of IE for project management. The team then traveled to Nepal and adapted the design to the capacity and needs on the ground. Throughout the IE process, the MoAD, MoHP, and WB

[5] Nepal is administratively organized into units of decreasing size: regions, districts, sub-districts (illakas), municipalities (VDCs), and wards. Nepal has 75 districts, each of which is divided into a number of VDCs, the number depending on the population size. There are 3,914 VDCs nationwide and every VDC has 9 wards.

teams are supported by a full-time DIME field coordinator based in Kathmandu who oversees day-to-day IE activities and ensures full communication across the various entities (research, operations and management).

The impact evaluation will affect policy through a few different pathways. The project operations team (at both the WB, MoAD, and MoHP) can use baseline data to help target project activities, and will use the IE results available at midterm to assess project effectiveness and make any necessary adjustments for the second half of implementation. The WB Country Management Units' (CMU) ownership of this evaluation will ensure that the resulting analysis impacts ongoing policy dialogue both at the country and sector levels. Brown-bag seminars will be organized between DIME and the CMU to ensure that the country office staff is aware of the learning coming out of this program, and that the IE work is timely and relevant for the CMU's decision-making process.

In addition to delivering evidence on key operational questions, this IE contributes to a large, global research agenda on aid effectiveness in agriculture (DIME-**aadapt**), which counts over 30 participating projects in over 20 countries. Representatives from MoAD and the Bank operations team have already participated in a capacity building and dissemination event of the **aadapt** community (Naivasha, 2012), and will share the results and experience from the ongoing IE work in future events, thus reaching a wide audience of policymakers worldwide

The Nepal IE team will produce high-quality research papers for presentation at research BBLs at the Bank (e.g. DECRG and DIME seminar series), events and trainings as well as international development conferences. The IE team will arrange dissemination events in Nepal, in which representatives from government, international organizations, and civil society will be invited to participate. These events will happen after each round of data collection. The findings will also be published as DIME policy briefs and submitted to peer-reviewed economics and field journals, thus reaching a wide audience of researchers and graduate students worldwide. All data will be made available online on the databank for IE, following the Bank's open data policy.

2. The Agriculture and Food Security Project

AFSP is designed to enhance the food and nutritional security of targeted communities through a holistic set of interventions. The project aims to improve the livelihoods of crop and livestock farmers, women engaged in household/kitchen-garden production, and households with pregnant and nursing women. AFSP will be implemented in 19 districts of the Mid- and Far-Western development region of Nepal,[6] and adapted to three agro-climactic zones: hills, low mountains, and high mountains. It is expected to reach about 162,000 beneficiaries.

[6]Darchula, Baitadi, Dadeldhuda, Humla, Jumla, Mugu, Dolpa, Kalikot, Bajhang, Bajura, Jajarkot, Achham, Doti, Dailekh, Surkhet, Rukum, Salyan, Rolpa, and Piuthan

AFSP takes a multi-sector approach to promoting agriculture and food security and is implemented jointly by the Ministry of Agricultural Development (MoAD) and the Ministry of Health Policy (MoHP). The MoAD is the lead ministry implementing AFSP. Project management and administration will be carried out by the PMU, headed by a Project Director deputed by the MoAD. The BCC intervention will be implemented by the Ministry of Health and Population (MoHP). The PMU will work closely with MoHP in the implementation of this intervention. In addition, the Joint Secretary from MoHP will be part of the Project Steering Committee (PSC), the main governing body for the project that will facilitate inter-agency cooperation.

The project consists of four complementary components:

1. Technology Development and Adaptation
2. Technology Dissemination and Adoption
3. Food and Nutrition Status Enhancement
4. Project Management

After consultations with the government and the World Bank teams, it was decided that the impact evaluation would concentrate on Components 2 and 3, which promote the adoption of more productive agricultural technology and better nutrition practices among target groups.

2.1 The Technology Dissemination and Adoption (TDA) component

The TDA component aims to introduce poor farmers to new production and management methods for both crops and livestock, to improve their yields and incomes. AFSP will promote improved seeds, higher input usage, and improved health of crops and livestock using a Farmer Field School (FFS) approach. Based on the principle of adult learning, the FFS approach organizes farmers into groups that meet to discuss challenges and work with extension agents to develop solutions. These groups become an important venue to promote and diffuse new technology. The TDA component also contains nutrition-sensitive interventions such as promoting backyard poultry and kitchen gardens.

The FFS approach builds on the Farmer Group (FG) approach to agricultural extension conventionally practiced in Nepal. In the FG approach, farmers are required to organize themselves into groups based on their interests or enterprise type. These groups can then access public extension workers who provide technical support, and conduct demonstrations for the group. Extension agents provide technical advice to farmers based on general recommendations, but are not trained to adapt their advice to the needs of specific farmers. Although agents undertake needs assessments with FGs, the results are rarely incorporated

into the planning of extension activities. Farmers are rarely provided with critical support services of inputs, credit and marketing.[7]

In contrast, FFS under AFSP will encourage farmers to interact with extension agents to develop customized solutions to their problems and adapt new technologies to their specific conditions. In addition to technical support and demonstrations, member farmers will be introduced to new inputs, which they will receive for free or at subsidized rates. In addition, extension agents will visit farmers in their fields on a regular basis to assess the implementation of the technologies and inputs demonstrated in the group.

2.1 The Food and Nutrition Status Enhancement (FNSE) component

The FNSE component works to improve both supply and demand for nutritious food through nutrition education and Behavior Change Communication (BCC).[8] The main goal of this component is to improve the nutrition of young children and pregnant or breastfeeding mothers.

FNSE will primarily be implemented through Mothers' Groups (which already exist in each ward[9]). Members will be encouraged to use the food made available through the agriculture interventions to improve and diversify their diets, and provided with nutritious recipes developed by the Department of Food Technology and Quality Control (DFTQC) using locally available nutritious foods, including those promoted by AFSP. Mothers groups will receive targeted agricultural interventions such as promotion of kitchen gardens and backyard poultry. These agricultural interventions will be delivered by the Ministry of Agricultural Development (MARD). Members of mother's groups receiving these nutrition-sensitive interventions will also likely be members of the crop and livestock FFSs.

In addition, Mothers' Group members will also receive Behavioral Change Communication (BCC) messages through the Female Community Health Volunteer (FCHV). The nutrition messages received through the mothers groups will be reinforced through home visits conducted by one or two additional members of the mothers group selected by the FCHVs. The focus of BCC will be on promoting early initiation of breastfeeding and improving the quantity and quality of complementary feeding. The BCC is designed to be complimentary to the agricultural initiatives of AFSP, including promotion of dietary diversity made possible by production of different types of food. The BCC messages will be developed by the project and

---

[7] Agricultural Extension Services Delivery System in Nepal. Food and Agriculture Organization of the United Nations, June 2010. ftp://ftp.fao.org/OSD/CPF/Country%20NMTPF/Nepal/AgExtServDelSysNepal.pdf

[8] This component also has sub-components on increasing food availability through food banks and decreasing women's drudgery, which will not be specifically studied as part of the evaluation.

[9] A ward is an administrative unit below a VDC. There are 9 wards per VDC.

several variations adapted to local social and gender contexts will also be produced. This intervention will target households with pregnant, lactating mothers and children less than 2 years of age to promote nutritious and diversified diets, appropriate infant and young child feeding (IYCF) and caring practices, and hygiene and sanitation practices. BCC will be delivered by the Ministry of Health Policy (MOHP), and will therefore be implemented on a different schedule than the agricultural interventions.

3. Impact Evaluation Motivation

3.1 Agricultural Outcomes

As explained above, the TDA component of AFSP uses a FFS approach to technology dissemination. In theory, FFS should provide improved results over FG through a number of mechanisms. One way to codify this is through Rogers' influential description of the "innovation- decision" process.[10] Rogers breaks down this decision into five steps: knowledge, persuasion, decision, implementation, and confirmation. While both FFS and FG 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 FG approach. Input subsidies make the initial adoption decision much easier. Continuous classes and instruction increase the likelihood of successful implementation of the technology, and thus also increases the likelihood of sustained adoption. While there are many different approaches to FFS, they all provide a far more intensive and prolonged effort to spur technology adoption.Given that FFS is more expensive than traditional extension approaches, it is generally executed with donor support and is not yet widespread. This evaluation will assist the GoN in understanding the comparative effectiveness of the intensive FFS approach in increasing agricultural productivity.

FFS will be the delivery mechanism for a variety of agricultural extension services including input distribution, training, and demonstrations. As all the agricultural interventions will be delivered as a package, this evaluation will not give a clean estimation of the effect of FFS versus FG with all else equal. Instead, it will measure the effect of AFSP's package of interventions delivered through the FFS approach, which will be compared to the status quo of FG extension services.

3.2 Nutrition Outcomes

The IE will test the effect on nutrition of AFSP's supply-side agricultural interventions, as well as the demand-side nutrition-specific interventions delivered as part of the FNSE component.

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

Education versus Agriculture to Increase Food Security

There is increasing agreement among development practitioners that food availability, while important, is not sufficient to eliminate malnutrition.[11,12] Rather, interventions to improve agricultural productivity need to become more nutrition sensitive, incorporating explicit consumption goals and measuring such outcomes. Pathways linking agriculture to nutrition are well articulated, but the evidence on the impact of agricultural interventions on nutrition outcomes is limited. A synthesis study of a wide range of agricultural interventions showed little impact on nutrition unless those projects made complementary human capital investments in nutrition education and gender issues.[13]A recent DFID-funded systematic review found that very few studies measuring the impact of agriculture interventions on nutrition used rigorous methods, and even fewer showed positive effects. The review determined that methodological weaknesses rendered the studies unable to detect positive effects, even if these did exist.[14]

Similarly, providing nutrition education (such as BCC) has mixed results on nutrition outcome. A recent Lancet meta-analysis[15] showed positive effects of education on breastfeeding, but no evidence of this leading to healthier infant growth. Education on complementary feeding strategies reduce stunting, but only when applied in food-secure environments or when combined with food or cash supplements.

We believe that this impact evaluation will make an important contribution in the literature by measuring the separate effects of nutrition-sensitive agricultural interventions and BCC on nutritional practices.

3.3 Improving the Effectiveness of BCC

The evaluation will delve into two key questions regarding nutrition education. The first is how relevant the BCC messages are to local gender and social contexts. The second is how those messages are actually delivered to mothers and the broader community.

---

[11] See for example Food and Agricultural Organization of the United Nations. http://www.fao.org/fileadmin/user_upload/agn/pdf/FAO_Approach_to_Nutrition_sensitive_agricultural_development.pdf

[12] See Anna Herforth, Andrew Jones and Per Pinstrup-Andersen Prioritizing nutrition in agriculture and rural development projects: Guiding principles for operational investments http://dyson.cornell.edu/faculty_sites/pinstrup/pdfs/wbdec2010.pdf

[13]Peter R Berti, Julia Krasevec and Sian FitzGerald (2004). A review of the effectiveness of agriculture interventions in improving nutrition outcomes. Public Health Nutrition, 7, pp 599-609. doi: 10.1079/PHN2003595.

[14]Masset et al. “Effectiveness of agricultural interventions that aim to improve nutritional status of children: systematic review”. British Medical Journal. 17 Jan 2012.

[15]Bhutta, Zulfiqar A., et al. "What works? Interventions for maternal and child undernutrition and survival." *The Lancet* 371.9610 (2008): 417-440.

BCC Messages Adapted to Local Gender and Social Contexts

The MoHP has a standard curriculum of BCC that is disseminated through its network of mother's groups. However, gender roles and low status hamper a mother's capacity to carry out infant and young child care necessary for higher nutritional outcomes even when she has the appropriate information. A mother's ability to breastfeed, introduce high quality complementary foods at the right time, or seek health services hinges on her ability to make decisions for herself and access resources such as time and money. In Nepal, behaviors are often determined by males in the household or elderly women. Only 46 percent of currently married women participate in decisions pertaining to their own health care, major household purchases, and/or visits to their family and relatives. Delays in seeking health care for easily-treated conditions, such as diarrhea, result in more serious and prolonged conditions among children and reduce their nutritional status. The high prevalence of domestic violence is one indication of women's low status in the household: 50% of women aged 40-49 report having experienced domestic violence at some point in their life.[16]

Social exclusion in Nepal along caste and ethnicity further compounds the problem of undernutrition and food security. Health services vary significantly across these social delineations. For instance, 40% of children are given oral rehydration salts during diarrhea as compared to 31 percent of Dalits, 38% of indigenous groups (janajatis) and 59% of hill Brahmans. The pattern is similar for percentage of deliveries in health facilities: it is 18% for all Nepalis but 70% for Terai Brahmans, 5% for Terai Dalits, and 14% for Janajatis. Media exposure, which is correlated with awareness of important public health information, also conforms to the pattern. For instance, 30% of Nepali women are not exposed to mass media even once a week as compared to 17% of males, 56% of Terai (or plains) Dalit women, and 31% of Terai Dalit men.

Given the importance of gender and social exclusion in Nepal, BCC programs will be more effective if the messages are adapted to the cultural context and gender dynamics in the household and community. This involves specifically targeting husbands, mothers-in-law, or other influential figures in the community when appropriate.[17] Addressing the high rates of malnutrition in Nepal requires improved IYCF practices by mothers which, in turn, require more active, strategically designed BCC messages. However, the development of tailored messages is time consuming and expensive. As such, the effort can only be justified if the resulting benefits in nutritional outcomes warrant the initial investments.

---

[16] Nepal Demographic and Health Survey Final Report, March 2012. http://www.measuredhs.com/publications/publication-fr257-dhs-final-reports.cfm

[17] Behavior change interventions and child nutritional status June 2011 Literature Review USAID/IYCN http://www.iycn.org/resource/behavior-change-interventions-and-child-nutritional-status-evidence-from-the-promotion-of-improved-complementary-feeding-practices/

BCC messaging adapted to local gender and social context are being developed by the Bank's Social Inclusion Action Program in Nepal, and will be piloted by AFSP. The NFSE Evaluation will test whether they are more effective than the standard BCC messages normally promoted by MoHP.

The relative impact of group and individual nutrition education

Existing evidence makes clear that in addition to the message formulation, message targeting and delivery is critical for messages to translate to behavior change. For example, one-on-one individual counseling is known to be generally more effective than group counseling.[18] The literature also shows that mothers and their families are more likely to accept and learn from a peer from the community. Finally, high intensity and frequency of message delivery is also shown to improve retention and application.[19]

In Nepal, Female Community Health Volunteers (FCHV) are expected to provide nutrition education to pregnant and lactating mothers through monthly mothers' group meetings and in-home nutrition counseling. However, given the increasing set of responsibilities entrusted to the FCHVs, home visits for nutrition counseling are generally omitted unless precipitated by an urgent need.[20]

In response to this reality, ASFP will encourage the FCHVs in selected VDCs to recruit two mothers from the mother's clubs to make weekly home visits to pregnant and lactating women in the ward. The objective of these volunteers is to reinforce the messages delivered by the FCHV during the monthly mother's club meeting, help mothers find solutions to individual issues, communicate with family members, and provide an accountability mechanism.

The IE will test the relative effectiveness of group education alone compared to group education combined with home based individual counseling by untrained mothers from the mothers group.

4. Evaluation Questions

1) What is the impact of AFSP's agricultural interventions on outcomes of interest? This includes the general agricultural interventions delivered by the TDA component, as well as the nutrition specific interventions of kitchen gardens and backyard poultry delivered by the FNSE component. Specifically outcomes are:

---

[18]Ibid, v.

[19]Inayati DA, etal. Improved nutrition knowledge and practice through intensive nutrition education: a study among caregivers of mildly wasted children on Nias Island, Indonesia. Food Nutr Bull. 2012 Jun; 33 (2): 117-27.

[20]http://www.measuredhs.com/pubs/pdf/FR181/FCHV_Nepal2007.pdf

a) Adoption of improved varieties, crop yields, and income for crop farmers
b) Livestock production and income for livestock farmers
c) Nutrition practices of pregnant mothers and young children

2) What is the impact of BCC on nutritional practices?
3) What the complementarity effect between BCC and AFSP's agricultural interventions?
4) To what extent can BCC messages adapted to local social and gender contexts enhance a mother's ability to apply the learning?
5) What is the impact of providing individual visits by peers to reinforce nutritional messages?

The IE will directly measure the causal impact of AFSP on two of the four PDO-level indicators defined in the results framework: increased productivity of crops/livestock, and improved dietary income for pregnant/nursing women and children.[21] It will also measure a number of intermediate and complementary indicators.

5. Evaluation Design

The main feature of the identification strategy is that the year AFSP operations begin in a given VDC will be randomly assigned. Additionally, the agriculture and BCC components will begin independently, allowing the evaluation to use a 2x2 design to estimate the effects of the components separately, as well as the interaction effects. In each district, AFSP will work in 10 VDCs chosen by local leadership, and project activities will phase in over the 5 years of the project.

5.1. Agricultural Evaluation

AFSP's agricultural components will be measured as a package, as they will all be delivered simultaneously by MoAD. This includes Farmer Field Schools for crops and livestock, as well as kitchen gardens and backyard poultry rearing, which will be delivered through mother's groups.

The evaluation of the agricultural interventions will use two techniques to establish a valid comparison group for those participating in AFSP. First, it will use a randomized phase-in of the VDCs over time: those selected to begin activities later serve as controls for the early starters. Second, VDCs that will not participate in AFSP will be selected as additional controls. Control VDCs will still have normal agricultural extension services (the Farmer Group system), so this evaluation will measure the effect of AFSP over the status quo.

Given the limited number of VDCs in each district, DIME and the PMU will jointly conduct a paired randomization of VDCs to ensure balance and improve power. The randomization process will proceed as follows:

[21] The two PDO-level indicators that will not be measured are: number of new technologies released for project-area farmers, and number of farmers with increased productivity. These indicators are more appropriately measured with monitoring.

1. Within each district, the 10 selected VDCs will be paired on the basis of geography, technical feasibility, food security and DAG score.[22] This exercise will result in five pairs per district.
2. 1 pair of VDCs per district will be eliminated from the impact evaluation. This pair of VDCs will start project activities in late 2014, which is roughly a year after farmer field schools begin in the first VDCs. These VDCs are eliminated because they will have limited exposure to AFSP before the follow-up survey.
3. From each remaining pair, one VDC will be randomly selected to be an "early starter" (FFS begins in 2013), leaving the other VDC to be a late-starter (FFS begins in 2015). This results in a total of 76 early starters and 76 late starters.

AFSP's agricultural activities will be phased in with early starter VDCs in 2013-14 and late starter VDCs beginning in 2015. Data for measuring impacts will come from a follow-up survey to be conducted directly before program operations begin in late starter VDCs. Differences in outcome indicators between the early and late starter VDCs at the time of follow-up will represent the causal effect of receiving AFSP for two years.

5.1.1 External controls

Using the randomly-selected "late starter" VDCs as the control group has the advantage that random selection provides a control group that is likely to best represent the counterfactual of the treatment population. However, as these VDCs have been selected for AFSP, there is some risk that they will receive project activities ahead of schedule, due to disbursement pressure, local politics, or even simply confusion. Also, local conditions may prevent random selection of VDCs in some districts. Therefore, we will also choose four VDCs per district not selected to participate in AFSP, to serve as additional controls. These VDCs will be matched to the "early starter" VDCs based on observable characteristics, using the same criteria as was used to create the pairs within project VDCs. These VDCs can also serve as long-term controls since they will never be part of AFSP.[23] The validity of these external controls will be established through balance tests of baseline data.

[22] The Disadvantaged Group (DAG) score is an index of indicators designed to provide a measure of livelihood at the VDC level.

[23] At the moment funding is only secured for two rounds of data collection, but long-term having additional controls allows for long-term follow-up if additional funding is secured.

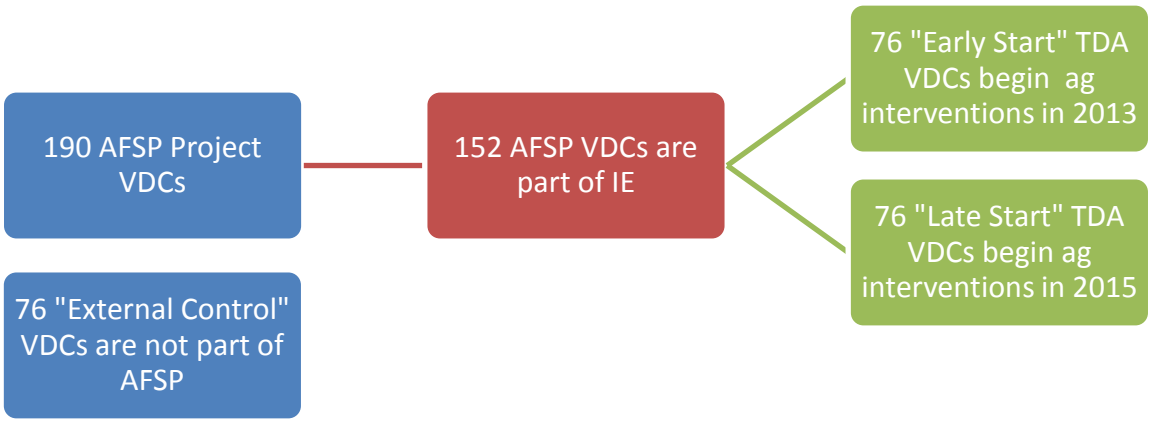

Table 1: AFSP VDC Phase-In

5.2. BCC Evaluation

As MoHP is coordinating the BCC component of AFSP, it will have a different phase-in schedule from the agricultural components. The IE will take advantage of the variable timing of these two components to identify the differential contributions of nutrition-sensitive agriculture versus nutrition-specific communication.

BCC will be phased in at the VDC level, in a manner analogous to the approach used for the agricultural interventions. Of the 8 VDCs that are part of the evaluation, 4 will be randomly chosen to receive BCC in the first year of their implementation (2014). 4 will receive BCC later, likely in 2016. The randomization will be stratified based on the treatment status for the agricultural interventions, and paired randomization will be conducted if there are large differences in observable characteristics likely to affect response to treatment, such as malnutrition rates. This process of orthogonal randomization ensures sufficient power for both the TDA and NFSE evaluations. Due to overlaps with other projects, BCC will only take place in 16 of the 19 AFSP project districts, so study of the effect of BCC will also be limited to 16 districts.

Table 1 shows the basic outline of the evaluation strategy for the two components.

	"Early Start" BCC VDCs begin BCC in 2013	"Late Start" BCC VDCs begin BCC in 2015	No BCC
"Early Start" Ag VDCs	32 VDCs	32 VDCs	12 VDCs

begin ag interventions in 2013			
"Late Start" Ag VDCs begin ag interventions in 2015	32 VDCs	32 VDCs	12 VDCs

Table 2: AFSP VDC Phase-In

In addition to testing these different approaches to improving nutrition, the IE will test variations in delivering BCC messages to beneficiaries. As discussed earlier, the two major treatments under discussion are:

- **BCC Messages Adapted to Local Gender and Social Contexts:** This intervention will have BCC messaging tailored to local contexts. The specific content of this additional tailoring is being developed by the Bank's Social Inclusion Action Program as a separate product.
- **Reinforcement of BCC Message from Peer Counselors:** In this intervention, BCC messages will be reinforced with home visits from other mothers in the ward. These visits serve to provide an alternate delivery mechanism for the messages and also increase their frequency.

Figure 2 shows the breakdown of VDCs into the various treatment arms. (For visual clarity, it does not show the treatments of adapted BCC messaging.) Out of the 64 "Early Start" BCC VDCs, half will be randomly selected to have the additional treatment of reinforcement from peer counselors. Additionally, half will be randomly selected to have the additional treatment of locally adapted BCC messages. The randomization of locally adapted BCC messages will be stratified based on the treatment status of the peer counselor intervention. The randomization assignments for the different interventions will be assigned orthogonally in order to maximize power of each individual test.

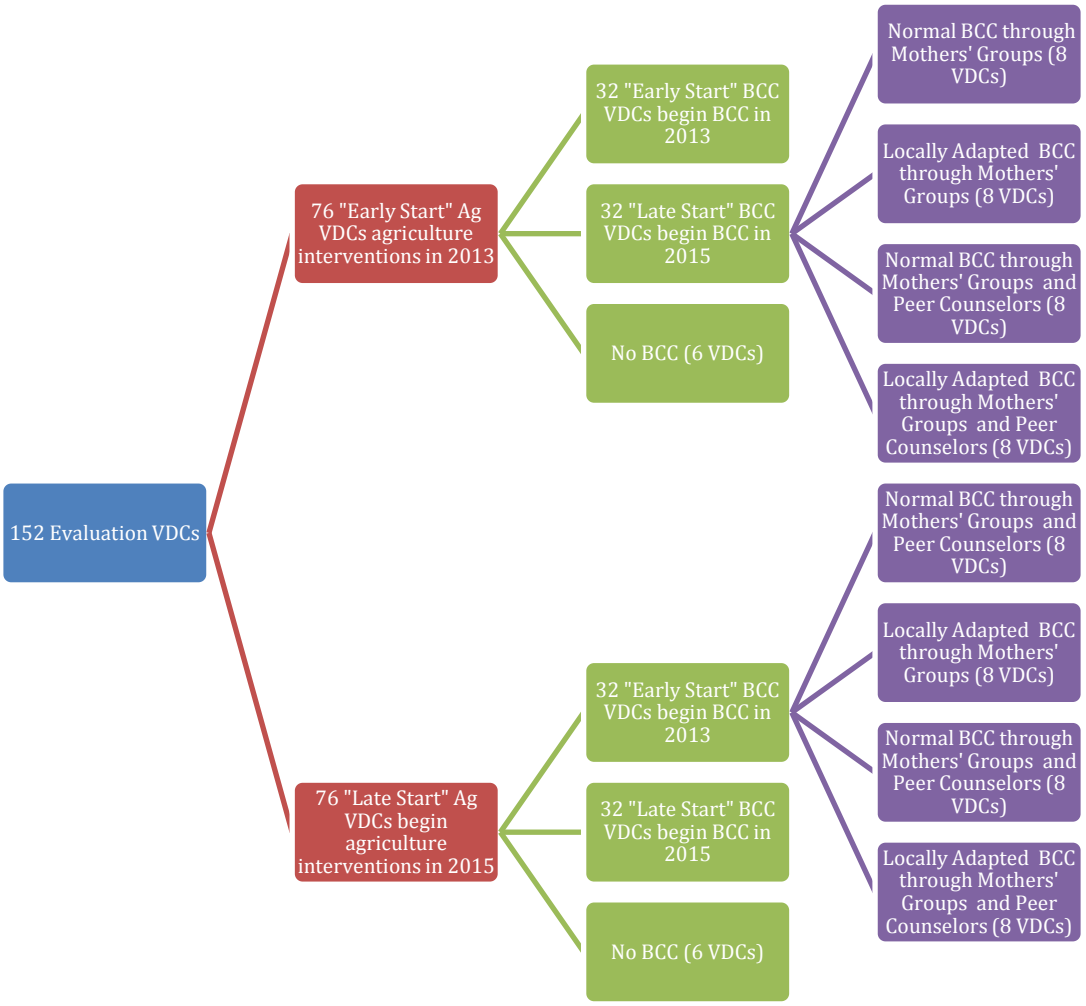

Figure 2: Evaluation Design

The breakdown of VDSs for the BCC interventions is clarified in Table 2.

64 "Early Start VDCs"	Peer Counselors	No Peer Counselors
Locally Adapted BCC	16 VDCs	16 VDCs
No Locally Adapted BCC	16 VDCs	16 VDCs

Table 2: AFSP VDC Phase-In

Note that there external control VDCs are not valid for the measurement of the effects of BCC. This is because in the majority of project districts, there are other programs (SUAAHARA and KISAN) performing very similar BCC campaigns in all VDCs. Therefore, our external controls will still have BCC.

6 Sampling

The IE sample will include will include 228 VDCs across 19 districts. This is 8 AFSP VDCs and 4 external control VDCs within each of the 19 districts. Within each VDC, data will be collected from 10 households at baseline. At follow-up, these 10 households will again be surveyed, along with an additional cohort of families with young children. This additional sample allows cohort analysis on indicators of childrens' nutrition.

6.1 VDC Selection

AFSP will be implemented in 10 VDCs in each project district. These VDCs will be selected by the District Agricultural Development Office (DADO) in consultation with local stakeholders. In selecting VDCs, DADO will strictly adhere to criteria (incorporating technical feasibility, disadvantaged group mapping[24] and food security data) outlined in AFSP's Project Implementation Plan (PIP). As mentioned before, the AFSP evaluation will be conducted in a randomly-selected 8 out of the 10 total VDCs.

Where possible, 4 additional VDCs that are not part of AFSP will be selected to serve as external controls. These 4 VDCs will be selected based on the AFSP eligibility criteria, excluding those that are receiving similar projects (such as the USAID-funded KISAN project).

6.2. Household selection

At baseline, the AFSP IE will sample 10 households from each of the 228 selected VDCs. The project will primarily affect households who are farmer group members. However, AFSP groups will not exist at baseline, and project activities start quickly after group formation. Hence, household selection will have to be completed prior to group formation.

A household census will be conducted in each project VDC to identify households that meet eligibility criteria for AFSP interventions, and thus compose the sample frame. The census will take place in wards where the first AFSP groups will be formed, in order to select AFSP participants who will have two full years of exposure to the project before the follow-up survey. The census will include a list of questions on household composition, land ownership and livestock ownership, and interest in participating in AFSP. The sample will favor households with young children who are most likely to benefit from all interventions: crop groups, livestock groups, and mothers' groups for nutrition.[25] The census and sampling will be coordinated

[24] The Disadvantaged Group (DAG) score is an index of indicators designed to provide a measure of livelihood at the VDC level.

[25] There is considerable uncertainty about whether households will commonly benefit from multiple interventions (specifically crops and livestock). If not, the sample will likely have to be increased at follow-up to ensure that enough members of various groups are included.

closely with the local officials who will be forming AFSP groups in order to ensure that the sampled households are very likely to end up joining the groups.[26]

At follow-up, the households sampled at baseline will be interviewed again, and additional households with young children will be added to the sample. Since the follow-up survey will take place two years after the baseline, households sampled for having young children may no longer have young children. Therefore, additional households will be added to conduct cohort analysis on nutrition indicators. The exact number will be determined using the baseline data

6.3. Power CalculationsPower calculations for agricultural outcomes were performed using agricultural yield data ($/ha) from the 2010-11 National Living Standards Survey (NLSS).[27] This variable has a relatively high intra-cluster correlation (ICC) of .17, which means that the amount of additional power from increasing the number of HHs sampled within a VDC is rapidly decreasing. This is the reason for sampling only 10 individuals per VDC as baseline. Due to uncertainty over whether households will benefit from multiple project components, for simplicity we assume that all sampled individuals will have the opportunity to benefit from all studied project components. In the event that this is not the case, we will increase sample size at follow-up to compensate, but for the purpose of power calculations we assume 10 households per VDC at follow-up.

Due to many treatment arms, different estimations will be made with different sub-samples of the data. Table 3 summarizes the main estimations, and the total number of relevant VDCs for each comparison. For the evaluation questions on just agriculture, the relevant sample size is 152 VDCs. For the nutrition outcomes, the sample is either 128 or 64 VDCs depending on which treatment arm one is exploring. For the purposes of these power calculations, we include only the AFSP VDCs and not the external controls.

Treatment Group	Comparison Group	Relevant Indicator	Total Number of VDCs
Ag Intervention early starters	Ag intervention "late starters"	Yields ($/ha)	152
BCC Intervention "early starters"	BCC Intervention "early starters"	3 IYCF Practices among infants 6-24 months (%)	128

[26] In some cases, AFSP groups will be formed out of pre-existing farmer groups. In these cases, our sample will come from these groups.

[27] Agricultural yields are the sum of value of sold crops plus estimate of value of self-consumed crops divided by area cultivated. Data is restricted to project districts. Nrs converted to USD at 85.6:1. Variable is winsorized at the top 1% tail to decrease the influence of outliers.

Early starters for Ag and BCC	Late Starters for Ag and BCC	3 IYCF Practices among infants 6-24 months (%)	64
Receive Locally Adapted BCC	Receive Normal BCC	3 IYCF Practices among infants 6-24 months (%)	64
Receive BCC + Peer Counselors	Receive Normal BCC	3 IYCF Practices among infants 6-24 months (%)	64

Table 3: Relevant Sample Sizes for Various Comparisons

We require a size of .05 and require a power of .85. Since we have no reference data on how well the baseline data will predict follow-up indicators, we estimate R at .3 for the agricultural indicators, which is a value seen in similar agricultural panels. However, we make the conservative estimate that the inclusion of the baseline data will not decrease the effective ICC in the impact regressions.

The estimation on yield will be taken using the full set of 152 evaluation villages, and the calculations yield a Minimum Detectable Effect Size (MDES) of .22, which corresponds to an increase of 26% in yields above the baseline value. The AFSP project documentation predicts an increase in yields of 30%, meaning that this goal should be detectable. Given the high within-village ICC, increasing the sample within VDCs adds very little to power, and the total amount of VDCs is constrained by the size of the project. Therefore, this level of power is that best that can be obtained given the constraints of the project. The results from the power calculations are shown in Table 4.

For nutrition outcomes, we consider the percentage of children age 6-23 months who adhere to the three standard Infant and Young Children Feeding practices (IYFC). We do not have individual-level data from our project districts on this indicator, but USAID reports[28] that 47% of Nepali children are fed according to three IYCF practices. We assume an ICC of .1, as this is a more standard finding for health indicators. Since this analysis will be a repeated cross-section, we make the conservative assumption that the R^2 is zero. For the estimation of the total effect of BCC (where 128 VDCs will be used), these assumptions give an MDES of .23, which corresponds to an increase 5.8 percentage points. Project documentation predicts an increase

[28]Infant and Young Child Feeding Update, USAID, Sept 2006
http://www.measuredhs.com/pubs/pdf/NUT1/NUT1.pdf

of 30% (14pp) of children between 6-24 months following three IYCF feeding practices. Therefore, our study appears to have enough power to detect the predicted effects.

The tests of variable BCC messages will take place only among the 64 "early starter" BCC VDCs. The test of interaction between BCC and the agricultural interventions will also have a relevant sample size of 64 (half of the early starter BCC VDCs and half of the late starter BCC VDCs.) With this sample, the evaluation has a MDES of .34, which corresponds to an 8.5 percentage point increase over the baseline value. We do not have a good estimate of the potential effects of our BCC variations or interaction effects, but this MDES could be interpreted as "large". However, the sample size is again limited by the number of project VDCs, making it difficult to increase power.

Indicator	Mean	Standard Deviation	ICC	R^2	Number of VDCs	Households per VDC	MDES	Change
Yields ($/ha)	458	539	.17	.3	152	10	.22	26%
3 IYCF Practices among infants 6-24 months (%)	47%	.25	.1	0	128	10	.23	5.8pp
3 IYCF Practices among infants 6-24 months (%)	47%	.25	.1	0	64	10	.34	8.5pp

Table 4: Power Calculations

7. Data

Data for the IE will come from two household surveys: a baseline survey conducted from August - October 2013 and a follow-up to be conducted from August-October 2015. It will collect detailed data on crop, livestock, and nutrition practices, and will directly measure 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 tons/ha for the four major crops supported by AFSP. The survey will also measure usage of inputs such as improved seed and fertilizer, along with adoption of promoted crops such

as vegetables. The collection of inputs will allow measurement of not only productivity but agricultural profits as well.

- Improved Livestock Productivity: The main indicator will be quantity of milk, meat, and eggs produced. It will also track the adoption of improved health and breeding practices.
- Improved Feeding Practices: This applies to pregnant and nursing women, and children between 6-24 months of age. We will measure these using indices such as the Household Hunger Score and Dietary Diversity Score. We will also ask questions about adherence to standard IYCF practices. Using different measurements will help us understand and identify the effects of the project on nutrition outcomes that are often multi-dimensional.

We will also include the following indicators

- Total Income
- Total income from agriculture (disaggregated by crop and livestock)
- Labor force participation
- Consumption
- Assets
- Housing

The Nepali non-profit research organization New Era has been contracted to conduct the baseline survey. New Era has over 40 years of experience doing social research in Nepal, including many household surveys for the Bank. The data will be used to establish a baseline for both the IE and for operations. (The project is not conducting a separate baseline survey.)

8. Internal and External ValidityThe main threat to internal validity of this study is that the assignment of treatment and control villages may not be respected. Although the PMU has agreed to the evaluation plan, local officials are not always aware of the need to avoid operating in control groups. Local officials might also face increased pressure from local political actors to implement some project activities in control VDCs. In order to mitigate this risk, the PMU has adopted a strategy of actively engaging district stakeholders in VDC selection and in impressing upon them the importance of the IE and respecting the VDC phase-in strategy. Additionally, the IE will sample 4 VDCs per districts that are not selected for AFSP to serve as additional controls.

Another threat comes from the presence of other development projects in the AFSP districts. In particular, the KISAN project implemented by USAID overlaps with 11 AFSP districts. This project also aims at improving agricultural productivity and enhancing nutrition and food security for its beneficiaries. If projects were to intervene in the same VDCs it would compromise the IE. To address this challenge, the AFSP PMU will work closely with KISAN's project team and to coordinate on VDC selection and ensure minimal overlap of project VDCs.

Since treatment and control VDCs are sometimes adjoining, there is a risk that spillovers could bias our treatment effects downward. We believe that spillover effects will be small due to the low population densities and difficulty in travel in the project area. However, exposure to the project by control VDCs will be closely monitored in surveys.

There is a worry that the study of the TDA component will not be as effective as possible since it will only measure the effects of the project in the first two years. Component 1 of the project (not part of the IE) is meant to test and release new agricultural technologies (primarily seeds), which would then be promoted as part of the TDA component. If the off-the-shelf technologies available to the project at the start are not effective, then the IE may underestimate the effectiveness of TDA.

Longer-term estimates of the project would be possible by adding an additional survey round, which would use the external controls as comparison groups. Additional funding would be necessary for this effort beyond what is provided by GAFSP. At the request of the CMU, the IE team will search for funding for an additional survey round.

This evaluation relies on close coordination between MoAD and MoHP. The two ministries will have to work closely to create an implementation plan that will allow for the evaluation of different approaches to improving nutrition. While the multi-sectoral nature of this project is one of the strongest elements of its design, it can also prove a potential risk if the ministries are unable to jointly implement AFSP's nutrition component. To address this, AFSP's Project Steering Committee (PSC) will include the Joint Secretary from MoHP as well as the Project Director from the PMU. This body will facilitate cooperation and coordination between the two ministries. In addition, the DIME Nepal Field Coordinator will meet regularly with officials from MoAD and MoHP to assist in developing implementation strategies for activities relevant to the IE.

Overall, the design is quite complex, and requires cooperation with a number of actors as well as continued progression of the project as planned. It is possible that some of the planned IE treatment arms may not end up being feasible. The implementation of the IE will depend on the progress of operations, and therefore the IE must adjust to the operational calendar. For instance, if development of context-specific BCC lags behing project implementation, this treatment arm may not be possible. The IE team plans to continually assess the feasibility of all aspects of the IE, and treatment arms will be dropped if they become impractical. These decisions will be made jointly with the operational team and the PMU.

The main threat to external validity is if the VDCs selected for our sample are not representative of Nepal as a whole. As VDC selection for the project will take place at the district level, it is not clear whether they constitute a representative sample. However, it is necessary for the evaluation to work within the constraints of the project, which can only work in areas with the required characteristics.

9. Evaluation Team

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

- Florence Kondylis, Senior 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
- Ritika D'Souza, Field Coordinator: AFSP Impact Evaluation Field Coordinator

The evaluation team also consists of two additional researchers, who have helped plan the evaluation design and will share in management of the evaluation:

- Slesh Shrestha, Assistant Professor, National University of Singapore
- Soham Sen, Consultant on Gender & Food and Nutrition Security, SASDS

10. Budget and Timeline

The impact evaluation is estimated to cost a total of $637420. These funds have been made available by GAFSP, and will not come out of project funds. A detailed budget can be found in Table 4.

Evaluation Design	Number	Units	Rate	Total
Field Coordinator (STC)	20	Days	280	5600
Social Specialist Time (Bank Staff)	7	Weeks	2200	15400
TTL Time (Bank Staff)	4	Weeks	2480	9920
Travel	6	Person-Missions	6000	36000
Baseline Survey				
Household Census	11400	Households	5	57000
Survey Costs	2280	Households	73	166440
Field Coordinator (STC)	90	Days	280	25200
Implementation Supervision				
Field Coordinator (STC)	90	Days	280	25200
Program Assistant (Local STC)	90	Days	100	9000
Follow-up Survey				
Survey Costs	2280	Households	73	166440
Field Coordinator (STC)	90	Days	280	25200
Report Creation and Dissemination				
Principal Investigator (STC)	30	Days	550	16500
Field Coordinator (STC)	20	Days	280	5600
Social Specialist Time (Bank Staff)	7	Weeks	2200	15400
TTL Time (Bank Staff)	4	Weeks	2480	9920
Travel	6	Person-Missions	6000	36000
Research Assistant (Baseline Data)	45	Days	280	12600
Total				637420

Table 5: Budget

The evaluation timeline for AFSP can be found in Table 5.

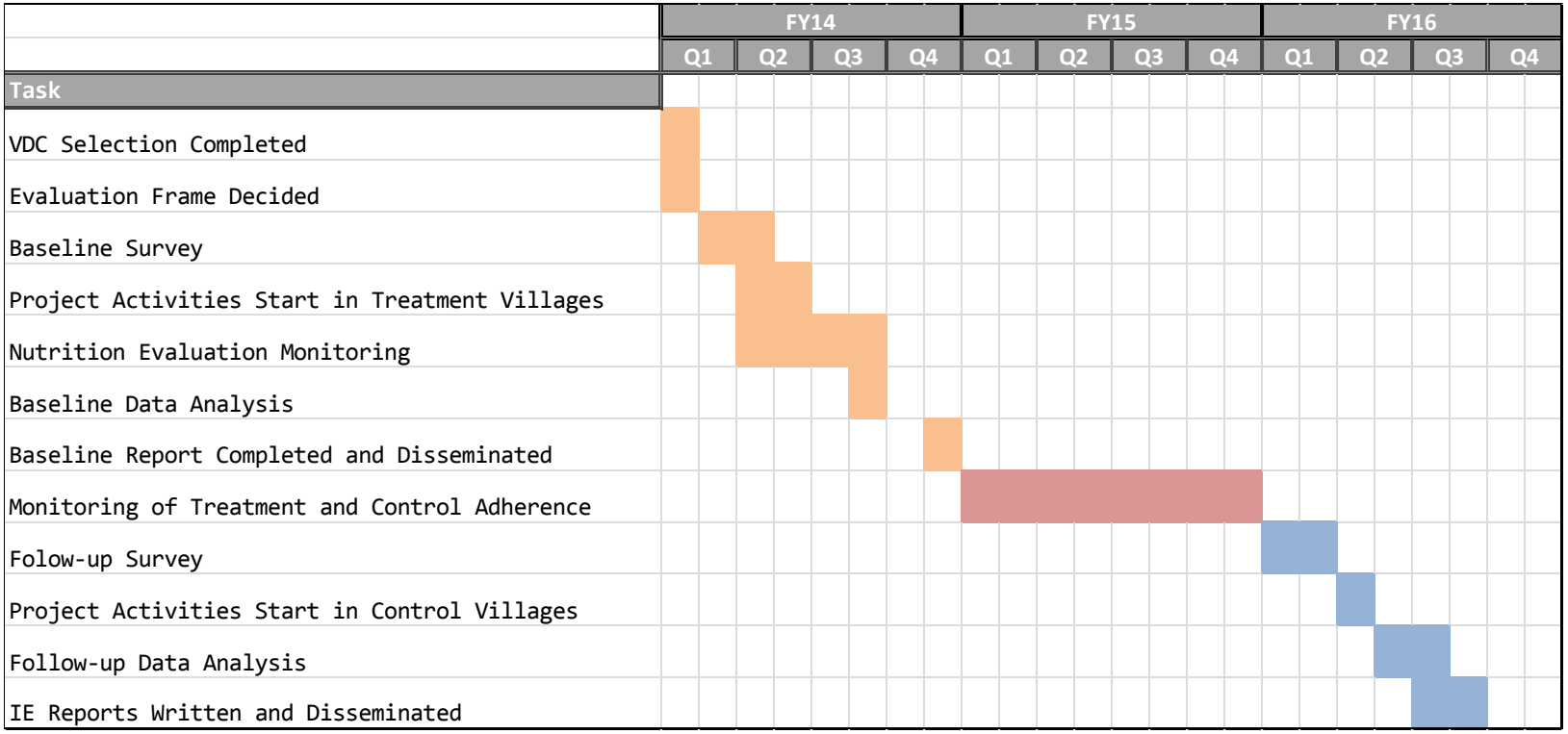

Table 6 : Evaluation Timeline