**SCOPE OF WORK**

**SAFE PROGRAM FINAL EVALUATOR’S TERMS OF REFERENCE**

**MERCY CORPS SUDAN**

| Programto be Evaluated: | Strengthening Agricultural Markets and Food Security (SAFE) Program |
| --- | --- |
| Donor: | Swedish International Development Agency (SIDA) |
| Location: | South Kordofan State (SKS) and Blue Nile Region (BNR) |
| Timeframe: | April – May 2023, for a duration not exceeding 35 working days |
| Contact: | | Robert Vokes  Director of Programs  rvokes@mercycorps.org | and | Sibongani Kayola  Country Director  skayola@mercycorps.org | | --- | --- | --- | |

# Overview

Mercy Corps, in partnership with the Adventist Development and Relief Agency (ADRA), is implementing the Strengthening Agricultural Markets and Food Security (SAFE) program in South Kordofan and Blue Nile states respectively. The program is funded by the Swedish International Development Agency (SIDA) and runs from July 2019 to end of April 2023, targeting 31,000 individuals. The main aim of the program is to reduce vulnerability and increase income and food security of male and female smallholder farmers, using a market systems development (MSD) approach.

During its inception, to understand the extent to which the market system was not functioning effectively for smallholder farmers, the team analyzed symptoms experienced by farmers, as well as underlying constraints at the wider market system-level. The following key constraints were highlighted:

* *Limited supply of agricultural inputs:* Input companies existed in both SKS and BNR at the state level but lacked capacity to reach small holder farmers (SHFs). In some localities, small input outlets existed with a limited geographical reach. Input suppliers were primarily selling to large farmers and did not see the value in seeking out a rural small holder farmer customer base.
* *Informal and Formal**Financial Services:* There were existing financial service providers with good capacity but incentives to work with SHFs was low as they were considered a high-risk demographic.
* *Farmer Coordination:* Farmer groups existed but were largely inactive with group capacity low and often formed only to access loans or formed by NGOs and government, to deliver inputs and training but often became inactive after the specific intervention. Farmer coordination was present but largely ineffective for SHFs versus larger farmers.
* *Post-Harvest Handling:* Storage facilities existed but did not have the incentives to expand reach into more rural areas due to low production of SHFs. Hermetic bag actors were absent while the government and NGOs were the primary providers of information around crop handling.
* *Output aggregation:* Output actors had moderate capacity and high incentives to build out rural buying networks while in BNR there was high capacity at state level and a strong network of buyers at the locality level but incentive to reach SHFs was low.
* *Climate and weather information services:* The Government of Sudan was providing weather and climate information broadcasted over the radio. One telecom company was also providing the information via a subscription-based service. Information was state or locality specific, which meant, it was not useful to inform decision-making at farm level.
* *Gender Norms*: Few women were present in agribusiness operations and most businesses did not view female farmers as a targeted customer-base, indicating gender norms play a strong role in shaping business incentives and inform the constraints women face in accessing markets, as well as employment.
* *Land tenure and rental agreements:* Land tenure was perceived as a sensitive and culturally charged issue with local leaders and government playing significant roles in the process. IDPs, refugees and some poor hosts rent plots via verbal agreements that are not fully binding.
* *Local natural resource governance:* Lands were mixed-use between cropland and pastureland; Livestock from both local households and pastoralists roamed freely, including on cropland, which created conflict. There was lack of fences or other ways of protecting or earmarking land boundaries beyond compounds and government buildings. Farmlands and pasturelands were often far away from local communities making it difficult to monitor and manage, and watersheds in SKS and BNS expanded beyond community boundaries.

The program therefore based on the above identified constraints, capacities, and incentives for actors, to design the SAFE program with the following program objectives.

**Objective 1:** Male and female smallholder farmers have increased and more climate-resilient yields.

**Objective 2:** Male and female smallholder farmers increase sustainable natural resource management practices

**Objective 3:** Male and female smallholder farmers have improved market linkages

**Objective 4:** Increased gender equity and increased empowerment of men, women, boys, and girls

**Objective 5:** Increased preparedness and ability to respond to disasters.

Since its inception in 2019, the program has experienced various changes, all aligned to better achievement of its goal of stimulating economic prosperity, food security and resilience by improving the performance of the agriculture sector in the two target states of South Kordofan and Blue Nile.

Mercy Corps will evaluate this program so as to learn from and document results, determine program impact, effectiveness and efficiency, examine potential sustainability of interventions and draw any lessons learned from the program. Results from this evaluation will also inform Market systems programming in Sudan and complex markets at large. This evaluation will be conducted in both states of implementation by a lead evaluator/evaluation team.

**Monitoring, Evaluation and Learning (MEL) within SAFE program**

SAFE program has an embedded Monitoring and Evaluation system led by the MEL Specialist with support from a MEL coordinator and MEL Officers. The key sectors in the program have results chains that are an internal planning tool, charting the logical flow of activities through to the desired impact and providing information on how specific interventions/activities fit into the overall strategic framework. The program’s monitoring plan was developed with a set of indicators, which were based on the results framework. The baseline survey and midterm evaluation collected both qualitative and quantitative information triangulating a synergy of various techniques (KII, FGDs, semi-structured interviews configured in smartphones) from a range of relevant stakeholders in the program areas.

**Scope of evaluation: Main Purpose and Objectives**

The main purpose of this assignment is to evaluate program achievements towards meeting the objectives and targets and to assess relevance, effectiveness, impact, efficiency and sustainability of the program at program end. The evaluation will also generate recommendations for future programming. The information will be used for public presentations and learning, for sharing with government counterparts, local stakeholders, and private sector partners, and for promotion of services in the community, as well as identifying possibilities for project replication. The results from the independent evaluation will also support Mercy Corps’ agency-level learning, by documenting and explaining why planned activities succeeded or failed. The final evaluation is expected to establish plausible links between program inputs and outcomes and results and draw lessons for improvement of future agricultural development and M4P (market facilitation) programs or similar future activities. More specifically, the evaluation should:

1. *Determine the extent to which program outcomes and their targets were achieved.*

The evaluator will be expected to analyze quantitative and qualitative data and report on the outcomes and results of the program on participant households. Outcome refers to the effect of the more immediate tangible benefits and includes access to services such as; inputs and finance, increased incomes, and improved performance of agribusinesses in the input and output (buyer) markets. The key program outcomes include;

* Outcome 1: Male and female smallholder farmers have increased and more climate-resilient yields
* Outcome 2: Male and female smallholder farmers increase sustainable natural resource management practices.
* Outcome 3: Male and female smallholder farmers have improved market linkages.
* Outcome 4: Increased gender equity and increased empowerment of men, women, boys and girls
* Outcome 5: Improved shock preparation and response

The evaluation will also capture the observable changes resulting directly from linkages between target smallholder farmers and the supporting agribusinesses and financial institutions facilitated by the program.

1. *Determine the effectiveness of programmatic activities and alternative implementation approaches.*

The evaluation is also expected to examine, as systematically and objectively as possible, how well the program attained its overall goal and strategic objectives and whether the achievements were efficient, relevant and sustainable. Focus may be placed on, but not limited to: what worked (why?) and what did not work well (why not?), provision of technical assistance (TA) through village based agribusinesses and market facilitation as an effective way of integrating vulnerable households into formal markets, relevance of the program and interventions to meet beneficiaries’ needs, and relevance of the program to government of Sudan strategic objectives, laws, and policies.

1. *Assess early indications of systemic change.*

A market systems approach has been employed to influence market actors to make new investments and adapt better and inclusive practices (models) that will in turn impact the lives of smallholder farmers in targeted areas. The evaluator will review the extent to which business models promoted by the program, have caused systemic change.

A change is defined as systemic, if it influences changes in commercial relationships and incentivizes structures or rules that govern access to products and services in a sustainable manner. The evaluator will be required to use the AAER framework, which is the latest Donor Committee for Enterprise Change (DCED) guidance on assessing system change. Using the framework, the evaluator will categorize changes related to how SAFE program partners and other local market actors have reacted to the following climate smart and inclusive business models promoted by the program;

1. Cooperative model
2. Village agent model
3. Contract farming model
4. Tree Nursery Operator model
5. *Examine the potential sustainability of established mechanisms and activities.*

This refers to how the program activities will continue after the program ends, such as the degree to which established market linkages and microenterprises are likely to continue without program support; community management committees will continue with viable operations; and beneficiaries will sustain relationships with the private sector (suppliers/distributors, buyers, producers). Overall, the evaluation needs to establish whether targeted beneficiaries will continue to have long-term positive benefits resulting from the program, including organizations whose capacities have been built by the program that may provide some continuation of the services once the activity has completed.

1. *Identify key lessons learned and best practices for replication.*

The evaluation is also expected to draw key lessons learned (positive and negative) in the past three years in implementing the program. The evaluator should illustrate best practices for replication in future programs. Areas of interest may include comparison of the market facilitation approach (crowding in potential market actors, etc.) that SAFE program has undertaken vs. the classic TA and distribution approach.

*Annex 2, located at the end of this Scope of Work, contains several key evaluation questions the consultant will be expected to address during the course of the evaluation.*

# Methodology

The evaluation will be implemented in both States of South Kordofan and Blue Nile region using a mixed methods evaluation methodology.

The quantitative household data has already been collected in both states and is being analyzed by the SAFE program MEL team using the same tool that was administered at baseline and midterm, with slight adjustments to capture modifications in performance management plan that happened mid-implementation. The SAFE program team has also conducted FGDs and KIIs with key stakeholders in both states and all data sets will be handed over to the consultant upon contracting. The consultant will therefore be responsible for aggregating all the available quantitative and qualitative information available and use it to compile the program’s final evaluation report reflecting overall conclusions and lessons.

Much as a lot of the data has already been collected by the SAFE program team, the evaluator will have the flexibility to include additional tools to allow for proper evaluation of the program. The consultant will also review previous assessment reports, progress reports, progress data, assessment data and depending on the need, will be supported to collect additional data, in case of any information gaps.

**Data Analysis and Results Dissemination Plan**

For quantitative data collected via the household survey, the SAFE program MEL team will conduct descriptive and bivariate analyses using the appropriate tests of statistical significance.  Where comparisons between groups (defined through bivariate analyses) are made, the level of statistical significance associated with the difference between groups will be cited.  Confidence intervals, p-values and other criteria used in hypothesis testing will be documented.

Qualitative information collected by the SAFE program team will also be synthesized with supporting quotations, to highlight key points and conclusions. Qualitative information will provide insights to quantitative results and is expected to uncover if, why and how program activities led to varying levels of outcomes as well as unintended and unexpected results.

The evaluator will present initial evaluation findings in a half-day debrief workshop attended by State leadership, donors, other NGOs, Private sector actors and SAFE consortium partners. Feedback from the workshop will guide the finalization of the second draft report that will be circulated to respective Technical support units of Mercy Corps and SIDA. Feedback from this will culminate in the final report. The consultant will be responsible for sharing of findings and lessons learned with Mercy Corps, and if and when requested by Mercy Corps, the consultant will share findings and lessons learned with the donor and government representatives. Hard and soft copies will be shared with SIDA and Mercy Corps for internal utilization.

# Evaluation Team

The consultant will be external, and will be responsible for compiling all the relevant program data and information from various program documents into the final evaluation report.

Mercy Corps global MEL team and technical support unit will provide input to the evaluation design and draft report. The consultant will report to the Director of Program and SAFE program consortium director, who together with the MEL specialist, will provide all the relevant program documentation (prepared or requested).

At minimum, the evaluator should have experience in market facilitation evaluation and programming, financial services, agriculture development, and food security in complex markets.

Mercy Corps MEL Specialist will support the team to ensure access to existing program data, reports and documentation, setting up meetings with key stakeholders if the need arises.

# Key Deliverables

The consultant will be responsible for:

* Comprehensive evaluation plan submitted within 4 days after contract award including;
  + Evaluation methodology
  + Suggested improvements to evaluation scope.
  + Revised evaluation timeline
* A presentation of initial findings to the SAFE program consortium team, Mercy Corps SMT and TSU upon completion of the initial draft evaluation report.
* An electronic version of the final presentation prepared based upon the initial findings of the evaluation. A summary version of findings should be shared as a final deliverable
* Draft and final evaluation reports as per specified format (see below).

Relevant documents to review include the Program inception report, Baseline and Midterm/Impact Assessment Report, sector specific internal assessment reports and all essential monitoring and evaluation tools and databases used by the program.

# Evaluation Timeline

The evaluation will be conducted in the months of **April and May 2023**, for a duration not exceeding 35 working days.

# Reporting and Dissemination Requirements

The final evaluation report will not exceed 40 pages, not including attachments. The report shall be structured in accordance with the following guidelines:

* **Cover Page (with a photo)**
* **List of Acronyms**
* **Table of Contents** that identifies page numbers for the major content areas of the report.
* **Executive Summary** (3 pages) should be a clear and concise stand-alone document that gives readers the essential contents of the evaluation report, previewing the main points to enable readers build a framework for understanding the detailed information within the report. It should highlight the major findings and recommendations as per evaluation objectives, including lessons learned and good practices.
* **Methodology**: It should describe the methodologies employed by the team during the evaluation including strengths and weaknesses of the methods with potential to alter conclusions reached, schedule of activities undertaken, description of any or special statistical analysis undertaken.
* **Results**: The results section should be structured based on the evaluation questions and objectives. It may for instance be helpful to organize the report against project objectives, while in some instances it may make more sense to organize the report and findings as per evaluation questions.
* **Synthesis, Recommendations and Lessons Learned**: This is space for the consultant to reflect on the data and results in a way that logically leads to concrete recommendations for current and future project improvements, pulls out organization lessons learned, and generally comments on data and results. Everything presented in this section must be directly linked back to the information presented in the results section of the report. Ideally, items discussed here should not be completely new to the reader, but rather will refer to previous discussions. Recommendations that are not directly tied to Results can be included in an Evaluator Comments section for the report.
* **Annexes**: These include information collection instruments in English and translation (if applicable); list of interviewees with numbers and types of interactions; SOW, qualitative protocols developed and used, photos, participant profiles or other special documentation needed. Datasets can be provided in electronic format and as an Annex.

# Required Expertise

The selected consultant is expected to have strong expertise in program evaluation, specifically, evaluations of market-based interventions, M4P (market facilitation) and agricultural development, as well as financial inclusion and services. The consultant must have experience in a variety of settings and working with a number of different people that include members of staff, government officials, local government extension officers, associations, private business actors and community members in rural and urban environments.

Successful applicants will have:

* Demonstrated experience conducting and managing all aspects of complex program evaluation
* Demonstrated understanding of market systems development/M4P approach and agricultural knowledge
* Experience in Sudan preferred.
* Advanced academic qualification in finance, business administration, economics, or economic development
* Familiarity with the Sudan agricultural context and market actors
* Excellent analytical, research, writing with concise and compelling communication skills; Strong visual presentation skills preferred.
* Language skills: Professional capacity in English

Applicants must submit the following documentation: technical and financial proposal that defines, among others, understanding of the terms of reference, methodologies, and costing of planned activities, description of similar assignments and references, curriculum vitae, and expected consultant fees on a per day basis and terms of payment.

Mercy Corps will examine all applications to ensure that they contain no amendments to the terms or any other irregularities and/or errors. To assist in the examination and selection process, Mercy Corps may, at its discretion, assemble and establish a Review Committee. Applications will be assessed according to the following criteria:

Technical Evaluation: The applications will first be evaluated on technical merits. The technical evaluation assesses the capacity of the applicant, based on its submitted technical documents, specifications, performance history, and references submitted. For this purpose, certificates and additional references may be requested and subsequently examined.

Financial/Cost Evaluation: The financial/cost evaluation is based on the cost of products and services presented in the application. The evaluation is intended to assess and confirm that applicants provide good value to Mercy Corps for the cost of all goods and/or services offered.

Other Evaluations: After ranking applications and applicants according to technical and financial criteria, Mercy Corps may consider other variables, including, but not limited to record of past performance, integrity, and social responsibility.

**Annex 1: Criteria to Ensure the Quality of the Evaluation Report**

* The evaluation report should represent a thoughtful, well-researched and well-organized effort to objectively evaluate what worked in the project, what did not and why.
* The Evaluation report shall address all evaluation questions included in the scope of work.
* The evaluation report should include the scope of work as an annex. All modifications to the scope of work, whether in technical requirements, evaluation questions, evaluation team composition, methodology or timeline need to be agreed in writing by the technical officer.
* Evaluation findings will assess outcomes on males and females.
* Limitations to the evaluation shall be disclosed in the report, with particular attention to the limitations associated with the evaluation methodology (selection bias, recall bias, unobservable differences between comparable groups, etc.).
* Evaluation findings should be presented as analyzed facts, evidence and data and not based on anecdotes, hearsay or the compilation of people’s opinions. Findings should be specific, concise and supported by strong quantitative or qualitative evidence.
* Sources of information need to be properly identified and listed in an annex.
* Recommendations need to be supported by a specific set of findings.
* Recommendations should be action-oriented, practical and specific, with defined responsibility for the action.

**Annex 2. Key Evaluation Questions**

1. *To determine the outcomes and results of the program.*

* To what extent has the program achieved its set objectives and targets?
* How have the program activities affected the direct and indirect beneficiaries’ income and food security?
* To what extent has program promoted technologies, models and practices been adopted?
* How have the program activities changed lives (more climate resilient yields, improved natural resource management practices, improved market linkages, increased empowerment for men, women, boys and girls, improved shock preparation and response) of households in targeted communities?
* Are there other unintended but important outcomes and results that have been realized in targeted communities as a result of program activities?
* To what extent has the program improved the capacity of local businesses which have been used as vehicles for delivering goods and services to targeted households?
* Do the stakeholders have a sense of ownership of the program? What are their views on program implementation and progress?

1. *To determine the effectiveness of programmatic activities and alternative implementation approaches.*

* What is the program status with respect to target outputs in terms of quantity, quality and timeliness? What factors impede or facilitate the production of such outputs?
* Does the monitoring and evaluation system appropriately address the program’s objectives and indicator targets?
* How effective was the technical assistance provided throughout the program? To what degree was the TA adopted among program participants?
* How effective was the program in targeting vulnerable households using its developed selection criteria?
* What aspects of the program were particularly ineffective?

1. *To determine the efficiency and relevance of the program*

* How efficient has the program been in attaining its goals and objectives? What has been the average cost per beneficiary taking into account pass-on activities, training through community extension trainers, etc.?
* Which interventions are most critical and/or effective in achieving project objectives and intermediate results? What lessons can be drawn from successful activities?
* How consistent are the program activities and outputs with the overall goal and the attainment of its objectives?
* Are the activities and outputs of the program consistent with the intended impacts and effects?

1. *To evaluate the extent to which the program’s cross-cutting themes have been addressed and integrated.*
   * How effective is the program at reaching women? What could be done to improve women’s participation?
   * What effect is the program having, if anything, on the livelihood of the women participants and their households?
   * How has the program affected the gender-based relationships in targeted households?
   * Are beneficiaries adopting desired practices or behaviors? Which practices have beneficiaries been more inclined to adopt, and why? Are there certain groups within the population with lower rates of adoption and why?
   * What effect is the program having, if anything, on the capacity of households to mitigate environmental effects of scaled-up production activities in their communities?
   * How can programs such as this one, improve and increase its impact on these crosscutting activities or others on beneficiaries and their households?
2. *To examine the potential sustainability of established mechanisms and activities.*

* What systems or activities have been put in place to ensure sustainability? What mechanisms have been developed to maintain the infrastructures created or rehabilitated after the program end?
* Are established market linkages, microenterprises likely to continue without program support? What capacities have program participants acquired, and how will they sustain relationships with the private sector (suppliers/distributors, buyers, distributors, producers)?
* Are program activities and technical assistance related to adoption of better practices sustainable, i.e., are participants likely to continue receiving TA after the program ends? Are businesses likely to continue operating and remain financially viable after the program ends? Are pass-on activities going to continue after the program ends?
* To what extent will target participants continue accessing long-term positive benefits after the program comes to an end?
* To what extent will other local or donor resources continue to be available to perform the activities the program now conducts that will require continuation after the end of the program?

1. *To assess early indications of systemic change*

* *To what extent have program partners adopted the program promoted business models?*
* *Have initial partner(s) continued to invest in adopted behavior without project support?*
* *To what extent have other market actors adopted comparable changes – either direct copies or variants of the original models?*
* *Have actors in the wider system adjusted and changed their behavior in response to the changes in the support functions and rules?*

1. *To identify key lessons learned and best practices for replication.*
   * What improvements can be made to the program design to improve results?
   * What improvements can be made in the implementation of the program in order to improve results?
   * What are the main lessons that can be drawn from the program experience since its inception?
   * In particular, what have been the main lessons learned regarding targeting and working with vulnerable households?
   * What corrective actions are recommended regarding the design, implementation, reporting, monitoring and evaluation of the program?
   * What actions are recommended to follow up or reinforce initial benefits from the program?