

Impact Evaluation of the Smallholder Horticulture Marketing Programme (SHoMaP)

Republic of Kenya

Draft approach paper

Contents

I. Background	4
II. Project Description	5
III. Preliminary evaluability assessment of the project	8
IV. Methodology	8
V. Organization and responsibilities	16
VI. Timeline	16

Annexes

1. References	18
2. Selectivity framework for the Impact Evaluation	19
3. Theory of Change (value chain)	21
4. Logical Framework	22
5. Evaluation Framework	24

Abbreviations and acronyms

IOE	Independent Office of Evaluation of IFAD
SHoMaP	Smallholder Horticulture Marketing Programme
IE	Impact Evaluation
M&E	Monitoring and Evaluation
RIMS	Results and Impact Management System
PSM	Propensity Score Matching

Impact Evaluation of the Smallholder Horticulture Marketing Programme (SHoMaP)

Republic of Kenya

Approach Paper

I. Background

1. In line with the IFAD Evaluation Policy and as decided by the Executive Board, the Independent Office of Evaluation of IFAD (IOE) undertakes one Impact Evaluation (IE) every year. Given their scope, the Impact Evaluations rely on extensive data collection and data analysis methods in order to gather attributable evidence on the effects of a project on its beneficiaries. In 2017-2018, IOE is undertaking its fifth impact evaluation. The programme selected for the impact evaluation is the Smallholder Horticulture Marketing Programme (SHoMaP) in Kenya. The project was selected using a comprehensive selectivity framework (Annex 2).
2. This Approach Paper lays out the terms of reference for the impact evaluation, including its methodological approach, scope and the process. The overall goal of the IE will be to assess whether the project worked or not, and why, and in doing so provide policy-relevant information and lessons learnt for the design of future projects and for country strategies. The following are its main objectives:
 - i) To measure, and in doing so, establish if the project interventions had an effect on individuals, households, and communities, and whether this effect can be attributed to the concerned interventions. To this end, an attempt will be made to evaluate all effects - positive or negative, direct or indirect, intended or unintended.
 - ii) To provide evidence for, and to be used as a critical input towards, the upcoming Kenya country strategy and programme evaluation.
3. The results of the evaluation are expected to contribute to better informed decision-making and learning about successful and unsuccessful approaches to increased incomes and reduced poverty and to promote greater accountability for the performance of future IFAD supported-projects, including those focussing on value chains. This IE will add to IFAD's database of impact evaluations and in doing so will strengthen IFAD's empirical knowledge of the agricultural and rural sector, one that is assimilated from the use of robust methodologies and based on attributable evidence.
4. In assessing the impact of the project on beneficiaries, the evaluation will assess key evaluation criteria included in the 2015 IFAD Evaluation Manual in order to provide a more holistic assessment of the project. These include the four impact domains under rural poverty impact criterion: (i) household income and assets; (ii) human and social capital and empowerment; (iii) food security and agricultural productivity; (iv) institutions and policies and criteria of relevance, effectiveness, efficiency and sustainability of benefits, gender equality and women's empowerment, innovation, scaling up, environment and natural resources management, adaptation to climate change, overall project achievement and performance of partners (IFAD and Government).

II. Project Description

5. The SHoMaP was designed as a six-year project to be implemented between 2007 and 2013. However, owing to the delays in completion of market infrastructure interventions the programme was granted a one year no cost extension and it completed in December 2014. Below is the description of the project for some selected indicators.

Table 1

Selected indicators for the project

Key Dates			
IFAD Approval	Signing	Effectiveness	Mid-Term Review
18-Apr-07	10-Jul-07	23-Nov-07	08-Apr-12
Original Completion	Actual Completion	Original Loan Closing	Actual Loan Closing
31-Dec-13	31-Dec-14	30-Jun-14	30-Jun-15
IFAD Financing			
Loan (USD million)	22.2	% disbursed	96
Grant (USD million)	0.5	% disbursed	100
Actual component-wise costs (USD '000) and share in total costs			
Component	Total	Share of total (%)	
A. Market Systems Analysis	260	0.8	
B. Institutional Strengthening	1 876	6	
C. Investment in Domestic Markets	23 810	76	
D. Programme Management	5 372	17.2	
Total	31 318	100	

6. The overall goal of the Programme was to reduce poverty among poor rural households by increasing incomes and reducing unemployment and underemployment in medium-high potential farming areas where horticultural production is an important source of livelihood by easing input and produce marketing constraints faced by small-scale farm households who produce horticultural crops primarily for the domestic market. Towards this end, the two Programme development goals were:
- i) To increase incomes and reduce poverty among poor rural households in medium-high potential farming areas for which horticulture is a source of livelihood;
 - ii) To increase the health and welfare of Kenyans by improving the quality and increasing the quantity of horticultural produce consumed within the country.

After the 2012 mid-term review (MTR) took place, the development goal was reformulated as follows: Contribute to reduced poverty and improved health among poor rural households in medium-high potential horticultural farming areas.

7. To achieve these development goals, the Programme's immediate objectives or purpose were:
 - i) To increase the output of and the net margins per unit of land earned by resource-poor smallholders from horticultural production for the domestic market;
 - ii) To increase employment opportunities in the production, processing and marketing of horticultural produce;
 - iii) To reduce the cost to consumers and increase the quality of horticulture products consumed domestically.

As for the development goals, the Programme's immediate objective was reformulated after the MTR as follows: Increased domestic horticulture productivity and improved functional input and produce marketing system.

8. The Programme sought to address inefficiencies and constraints in input supply and horticultural marketing in target areas with the ultimate aim of:
 - i) Reducing into-farm unit cost of inputs among smallholder horticultural farmers;
 - ii) Improving the quality of inputs and services provided by input suppliers (stockists) to smallholder horticultural farmers;
 - iii) Raising the quality of horticultural produce traded in the domestic market;
 - iv) Increasing and stabilizing farm-gate prices for smallholder horticultural producers
9. **Target Beneficiaries, Geographical and Commodity Coverage.** Direct target beneficiaries of the Programme included smallholder horticultural farmers primarily producing for the domestic market; produce traders; input suppliers (stockists); produce transporters; horticultural processors; while indirect target beneficiaries include horticultural consumers; rural underemployed and unemployed men and women. Overall, the Programme benefited an estimated 60,000 farmers (12,000 farm facilities/households) through some 530 farmer groups. The outreach of the project (number of beneficiaries reached) by type of intervention can be found in Annex 5 of the Project Completion Report.
10. As per Programme design, the target areas comprised eight (8) horticultural producing districts spread across five provinces. The districts were selected using district targeting procedures developed during Programme formulation based on a weighted set of indicators relating to poverty profile, horticultural production, and the presence of other relevant initiatives such as the Kenya Agricultural Productivity Programme (KAPP) and the National Agricultural and Livestock Extension Programme (NALEP) for effective synergy. However, by the time the Programme started, the eight (8) original target districts had further been subdivided into fourteen (14) districts which were maintained as the administrative units of the Programme comprising Bungoma North, East, South and West in Bungoma County; Nandi South in Nandi County; Bureti district in Bomet County; Kisii and Gucha districts in Kisii County; Nyandarua North and South districts in Nyandarua County; Embu district in Embu County; Meru Central, Imenti South and North districts in Meru County.
11. The Programme targeted three horticultural crops per district which were selected through a participatory process involving all relevant stakeholders - an important tenet of the Programme at all levels of implementation. The commodities which had been identified by stakeholders as having the greatest potential for income

generation and poverty reduction include bananas (12 districts); tomato (9 districts); Irish potatoes (4 districts); onions (4 districts); mango (3 districts); passion fruits (3 districts); cabbage (2 districts); local/traditional vegetables (2) garden peas (2 districts); pineapple (1 district).

12. **Components.** The four components of the projects were:
- i. Domestic market system analysis. The main objective of this component was to provide a sound basis for targeting and prioritising Programme activities and in this respect inform decision-making in relation to the improvement of input and output markets and related interventions. Towards this end, the Programme was expected to carry out the following set of studies to inform the design of Programme interventions: i) three priority horticultural commodities in each of the 14 Programme districts; ii) two nation-wide studies: an Upstream inputs supply systems study; and a Downstream horticultural produce wholesale and retail marketing study; iii) 14 district-based value chain analysis (VCA) studies (one in each of the Programme districts) covering the three select crops; and iv) a district-wide stockists mapping study;
 - ii. Institutional strengthening. The main objective of this component was to support capacity strengthening needs of both service providers (including PMU, Government staff of collaborating ministries and other service providers including input stockists, brokers, wholesale and retail traders) and farmers and farmer groups. It had five sub-components: (a) support to existing formal and informal farmer groups through training in improved agronomy practices and marketing; (b) training of horticultural input stockists, traders, brokers and market managers; (c) on-the-job training of MoA staff; (d) support to evolving systems that provide market information to farmers and traders; (e) support to the development of improved horticultural sub-sector policy and legislation framework.
 - iii. Investment in support of domestic horticulture value chains. This component aimed to support the cost-effective investments and innovative initiatives that would have the real potential to break constraints facing the domestic horticulture value chain, add value to produce, reduce marketing costs, and enhance efficiency and equity with which marketing chains moved commodities from farms to markets. The above activities were to be supported by means of three sub-components: (a) pilot investments through competitive grants; (b) spot improvement of rural roads to provide accessibility and (c) development of downstream physical market infrastructure. The component was implemented through collaboration with Ministries of Roads, Public Works and County Government.
 - iv. Programme management and coordination. The State Department of Agriculture was responsible for overall implementation, supported by a Programme Management Unit (PMU) that facilitated and coordinated Programme activities in the programme area. The PMU was also responsible for financial oversight, monitoring, impact assessment and evaluation.
13. The Programme used participatory value chain development approach whereby systemic constraints and interventions were identified through stakeholder consultations¹ with the implementation based on the principle of seeking to strengthen the natural development of value chains rather than promoting radical alternative forms of structure, organization and trading practices. The Ministry of Agriculture (MoA) was the lead implementing agency while collaborating agencies included the Ministry of Public Works (MoPW), the Ministry of Roads (MoRs), the

¹ Stakeholders includes all persons associated with the project, including private sector.

Ministry of Local Government (MoLG) and other relevant ministries and wide range of stakeholders including farmers and farmer groups.

14. The Programme Management Unit (PMU) - a semi-autonomous unit located in Nakuru - was charged with the responsibilities of coordinating and facilitating implementation of Programme activities. The unit undertook these functions through district implementing units which comprised district and divisional staff of MoA and collaborating ministries; other relevant Ministries and stakeholders as appropriately identified at these levels. At the MOA, headquarters, the Programme supported the activities of the Agricultural Sector Coordination Unit (ASCU) for the development of the National Horticulture Policy. At the higher (policy) level, the Programme is guided by a Programme Steering Committee (PSC) which was chaired by the Permanent Secretary-MOA.

III. Preliminary evaluability assessment of the project

15. In addition to the selectivity framework that assists in selecting a project for the IE, an evaluability assessment is undertaken with an aim to give priority to projects that have an adequate amount of usable self-evaluation data to ensure that impact evaluations by IOE can be done in an effective and efficient manner. Availability of data helps reduce the costs and time taken for IOE to undertake impact evaluations.
16. The list of all sub-counties where the project was implemented is available, and the list of all commercial villages (villages where producer groups were trained) is also available. In addition, annual monitoring reports (in terms of outputs achieved), M&E and Value Chains analysis reports are available.
17. The project undertook both baseline and endline studies. However, the baseline study was prepared in 2011, late into project implementation (during the mid-point year of project implementation). Furthermore, it was done as one baseline in each of the 14 sub-counties of project area. Thus, there were 14 separate studies instead of one comprehensive baseline. Also, a fixed number of 150 respondents were sampled in each sub-county rather than having a proportional sample. The baseline study was conducted using only beneficiaries as respondents.
18. The project conducted an impact assessment at the time of project completion; it was conducted using quasi-experimental method with a comparison group using mixed methods. A total sample of 2,852 households, out of the total estimated 12,000 households, was interviewed. This included 2,187 beneficiaries and 665 non-beneficiaries for comparison group. The recall method was used to construct some of the baseline indicator values. However, the formal method used for selecting the comparison group which is a key requirement for establishing internal validity and therefore for attributing project effects is missing in the methodology.² In addition, the size of control group is far lower than the beneficiary group. Consequently, the degree of covariate balance achieved between the beneficiary and comparison groups also cannot be established.
19. Finally, at the time of its conducting, some of the project activities viz., physical market structures were still not completed, and hence the expected impact of the project could not be assessed.

IV. Methodology

The evaluation approach

^{2 2} It can be surmised through interviews with staff involved in the survey that control areas were selected based on geographic proximity to the treatment areas.

20. In light of the results of the above preliminary evaluability assessment, an important aim of the proposed IE exercise will be to design and implement a robust methodology that overcomes the inherent shortcomings in the available data and improves the validity of findings. But importantly, the exercise will be more comprehensive and cover several more areas (IOE criteria). In order to do so, the impact evaluation will use established methods for matching comparison and treatment units (such as the Propensity Score Method), appropriate sample sizes for the above two types of units and using probability proportionate to size (PPS). This section of the report will elaborate the methodology to be employed.
21. As mentioned earlier, in addition to evaluating the four impact domains (under the rural poverty impact criterion) namely, : (i) household income and assets; (ii) human and social capital and empowerment; (iii) food security and agricultural productivity; (iv) institutions and policies, the IE will assess other criteria such as relevance, effectiveness, efficiency and sustainability of benefits, gender equality and women’s empowerment, innovation and scaling up, environment and natural resources management, adaptation to climate change, overall project achievement and performance of partners (IFAD and Government).³
22. In line with the IOE Evaluation Manual, the above criteria will be rated on a scale from 1 to 6, with 6 representing the best and 1 the worst score. Moreover, project ratings falling into the three higher ratings (4-6) will be classified as “satisfactory” while the three lower ratings (1-3) as “unsatisfactory”. The results of the evaluation will inform the overarching impact evaluation report, which will be prepared by IOE, once the impact survey data and analysis are available.
23. The impact evaluation will be undertaken by a firm specialised in quantitative and qualitative methods, and which will be selected using a competitive process by IOE. In addition, a team comprising IOE staff and consultants will also undertake a mission to the country and will evaluate the project using methods such as observation, interviews and group discussions.
24. **Evaluation questions.** The impact evaluation will assess the successful impact of the project, or the lack thereof, by seeking to answer a set of questions at both overall project level and at specific levels; for the latter, the evaluation will probe the different points of the value chain where interventions occurred, both upstream and downstream, and in doing so, will attempt to establish which elements of the value chain were instrumental in creating the desired impacts and which were not:⁴
 - Did the project contribute to better socio-economic condition (incomes, assets and food security) of project beneficiaries, including of women beneficiaries (and youth beneficiaries, if possible), after project completion as compared to that of non-beneficiaries? If so, to what extent?
 - Did agricultural productivity of farmer beneficiaries increase as a result of project interventions, and if so, to what extent? What explained the results:
 - Support provided to farm input stockists resulting in better quality farm inputs and/or lower costs to buyer farmers?

³ The IOE Evaluation Manual 2015 provides details of these criteria. It can be accessed at: <https://www.ifad.org/documents/10180/bfec198c-62fd-46ff-abae-285d0e0709d6>

⁴ The list is indicative and may be modified during the planning of the impact evaluation.

⁵ These are a set of specific questions only. Questions related to all criteria are indicated in the Evaluation Framework in the Annex.

- Support to farmers through agronomy services provided by project-trained extension staff resulting in more suitable agricultural practices?
- Support to local seed producers especially for potatoes to improved local availability of quality planting material
- Did the training of farmer groups in marketing related activities and their linkages with private sector (buyers) lead to better farm gate prices through bulk selling of produce and sensitization of buyers?
- Did the construction of rural roads/bridges result in reduced transportation costs and post-harvest losses?
- Did the construction of physical markets, including storage spaces, result in better prices for consumers and reduction of losses to traders?
- Did product value-added through the pilot initiatives lead to income diversification for the beneficiaries?
- Did women's empowerment increase as a result of the project?
- What other (unexpected) impacts (positive or negative) did the intervention have on the wider community?
- Overall, what were the key factors that enabled project impact? What factors, including exogenous factors such as the devolution to counties, hindered the impact?

Theory of change of the project

25. The theory of change is the point of departure for this IE (displayed in Annex 3). It demonstrates the causal pathway from outputs to outcomes (short and medium to long term) and finally to impact. Whilst the theory of change is also an extended expression of the log frame (see Annex 4 for log frame), the one to be considered for this project is reconstructed. In other words, it takes into account some of the main changes that occurred during the project implementation, especially with regards to activities and outputs. To this extent, it differs from the log frame that was developed at the appraisal stage and which was not modified to reflect the changes as they occurred.
26. Importantly, the theory of change is cast in a value chain structure, which is essentially the underlying premise of the project. Thus, it shows both vertical and horizontal linkages, the former indicate forward and backward linkages between upstream and downstream actors resulting from project interventions, and the latter indicate how activities and outputs related to the same actor result into outcomes (for that actor). As depicted by the figure, the interventions lead to common medium-long term outcomes such as increased value of production and improved food security. The impact or the goal of the project is an increase in the incomes. One more objective of the intervention logic is to present the assumptions that underpin the transition along the causal path (this is shown by way of coloured boxes).
27. In terms of methodology, the various links in the intervention logic will be analysed using a variety of methods, building up an argument as to whether or not, and to what extent, impact pathways have been realized in practice. This will essentially help answer the "why" question i.e. why interventions have or have not worked and will thus complement the findings on impact.

28. In the context of a value chain project, it will be important for the evaluation to demonstrate the validity of the relationships or the forward and the backward linkages between the various actors in the chain.

Mixed methods approach

29. The IE will use a mix of both quantitative and qualitative methods in order to utilise the strengths, and overcome the shortcomings, of each of the two. Although scant evidence was found of the use of mixed methods in value chain impact evaluations, combining qualitative and quantitative approaches to data collection and analysis is preferable (ILRI 2014). The two methods will be undertaken in parallel, for reasons of cost and time efficiency. These methods will be used to answer the key evaluation questions for the evaluation criteria and which form part of the evaluation framework. The evaluation framework is presented in Annex 5. A table outlining the choice of method to evaluate the four impact criteria is provided below (Table 2).

Table 2

Choice of methods for analysis of selected variables under the impact domains

Impact Domain	Focus of Inquiry	
	Quantitative	Qualitative
Household income and net assets	Household incomes/ expenditures. Resilience (diversified sources of income). Household assets, productive and others. Use of income.	Effects of shocks and disasters
Human and social capital and empowerment	For women-headed households (increased income and access to assets and infrastructure, food security).	Gender relations (decision making: intra- and extra-household; workload distribution). Community relations; acquisition of new skills and knowledge; Social capital (e.g. relations with traders, service providers etc.) (Venn diagram)
Food security and agricultural productivity	Household food security and nutrition (increased intake/dietary diversity). Average cost/value of production, average area under production, average yields per hectare, average value/volume of post-harvest losses, average income from value added agriculture.	Food availability (seasonal mapping)
Institutions and policies		Influence on policies/practices of project interventions

30. Since the assessment is ex-post, the impact assessment will use quasi-experimental design in order to address the issue of endogeneity bias and to better attribute project results to the project interventions. Any identification of impact will be achieved through a counterfactual, i.e. what would have happened to a treatment group in the absence of the treatment. Quasi-experimental evaluation frameworks are more adaptable to value chain approaches (ILRI 2014).

Quantitative research design

Household Survey

31. In order to measure impact, a household survey will be carried out in treated and comparison communities, allowing assessment of project interventions using the with/without (treated/control) approach. The unit of analysis will be farmer groups who received training from the project. The household questionnaire will be designed by IOE in-house. It will consist of questions regarding general household characteristics, such as household demographic characteristics, land use, on-farm and off-farm employment, income and asset ownership, among others. Since the baseline study was not conducted before or at the time of project start, the survey will recreate the baseline using recall data for the starting year of the project , avoiding its use for variables such as income and expenditure where recall data can be heavily influenced by response bias. Options to include secondary sources of data, such as Kenya's the Household Budget Survey that was undertaken in 2006, will be explored for recreating the baseline values.
32. Unlike randomized experiments, where the likelihood of being selected is the same for the treatment group and the control group, and their distributions of observable and unobservable characteristics are equivalent in a statistical sense, nonrandomized selections are likely to suffer from biases (IEG, 2010). Of the several possible sources of bias, two are relevant to this evaluation: non-random project placement and non-obligatory project participation.
 - *Non-random project placement:* this refers to the fact that the project was implemented only in areas with a larger horticulture production. Thus, areas that received project support in this regard are likely to differ from areas that did not.
 - *Non obligatory project participation:* this refers to the training for farmers, stockists, traders and transporters, in which some persons participated and some did not. Those who chose to participate could differ from those who did not.
33. These biases can result in overestimates or underestimates of the impact because the two groups in each case may not have been statistically equivalent before the project started. In order to overcome these selection biases, propensity score matching (PSM) will be undertaken wherein the calculation of propensity scores⁶ is based on characteristics discussed below. The propensity score or conditional probability of participation will be calculated by using a probit or a logit model in which the dependent variable will be a dummy variable equal to one if the farmer participated in the project and zero otherwise. The vector of independent variables will comprise those characteristics that determined project placement. The variables included as independent variables are those that will be exogenous to project participation i.e. not affected by participation in the project.
34. Matching will be done at the sub-county level and then at the cluster (village) level. This process will lead to the identification of villages which will be good comparisons to the treated communities. Propensity scores for each village will be calculated using the following characteristics from baseline year.
 - Population size
 - Gender composition
 - Distance to the market
 - Distance to a significant road
 - Predominant settlement type (% rural)

⁶ Propensity score matching (PSM) is a statistical matching technique that produces a comparison group similar to the treatment group with respect to measured characteristics, controlling for observed differences prior to treatment and enabling an "apples to apples" comparison.

- Monthly need-based subsistence subsidies (per capita, family, registered needy family)
 - Areas with horticulture potential
35. The selection of areas for control groups will take into account the type of horticulture commodities that are produced; these will be the same as those produced in the treatment areas. In order to avoid spill over effects, care will be taken to ensure that control group areas are not serviced by the traders and input suppliers who were trained by the project.
 36. To further minimize biases from differences treatment and control groups, the analysis will use an estimation approach that compares the average changes between the former and the latter before and after treatment. This difference-in-differences estimation approach controls for time-invariant fixed effects by differencing them out of the estimation; this procedure though assumes that the treatment and control groups are on statistically similar trends.
 37. Sampling individuals. Interviews at the household level will be performed in the selected clusters (both treatment and control clusters). To sample individuals, the random walk method will be used. Use will be made of Geographic Information System to identify starting points for random walk and to make sampling more efficient by avoiding non-inhabited structures.⁷
 38. Following treatment community fieldwork, data analysis will be carried out to develop a profile of each type of beneficiary. Based on these profiles, questions will be developed for a screener questionnaire for control communities (tagging questions). The screener questionnaire will identify whether a household in a control community would have been likely to be a beneficiary of the project had it been implemented there.

Sample size

39. Any survey aimed at assessing impact must start with power calculations and a rough understanding of the minimum effect size to be detected in order to sample an appropriate number of individuals. In order to calculate the sample size necessary to pick up this effect size, standard power calculations will be performed.
40. The sample size will be calculated using the following parameter values: alpha=0.05, beta=0.2, a Minimum Detectable Effect of 0.20 for income variable (assumption based on the project endline survey), an intra-cluster correlation value of 0.1 and adjusting for possible non-response (5 per cent), a sample size of **1500 households** will be obtained, with 700 in the treated group and 800 in the control group. The oversampling of the control group is in order to find the best quality matches possible for the treated group and to confront the issue of the control group sampling units dropping out due to lack of adequate matching. Probability proportionate to size (PPS) approach will be used to distribute sampling units across project locations.

Survey instrument

41. The survey instrument will contain question blocks that will enable to assess project impact on household income and net assets,⁸ human and social capital and empowerment, food security and agricultural productivity. In addition to these question blocks, the questionnaire will contain a standard household passport section, including demographic questions for household members and a number of questions attempting to understand of the various interventions.

⁷ The sampling frame will be created from records with the M&E unit that has the names of villages and of households who are group members. Where these records do not exist, and for the control group, mapping of sampling units will be undertaken with the help of local officials and village heads.

⁸ Insofar as possible, the focus will be on horticulture income.

42. Possible measurement errors arising from recall data are possible and hence attempt will be made to use variables that are easier to recall (changes in variables such as income, profit, and so forth are difficult for subjects to remember while major asset purchases are the kinds of variables best used within this framework) and to use major events around the recall period, if these exist. The possibility of using secondary data will be also be explored for baseline values.

Fieldwork

43. Pilot interviews. Once the questionnaire has been developed, pilot interviews will be carried out. The interview reports will focus on how respondents understood questions and whether any question seemed difficult for respondents to understand. If any issues are uncovered with the questionnaire during the pilot, cognitive interviews will be carried out to understand where issues are and to test alternative question wordings.⁹
44. Fieldwork supervisor and enumerator recruitment and training. Enumerators will be recruited based on their experience working on similar projects and enumerator home region. These selection criteria will yield high-quality results through speedy fieldwork through the use of enumerators already in the proximity of the selected communities.
45. Fieldwork, back-checking and data cleaning. Following interviewer training, fieldwork will be carried out. Enumerators will interview respondents using face-to-face computer assisted personal interviewing (CAPI) on tablet computers. This form of data collection allows for greater accuracy, increased speed of fieldwork, enhanced quality control mechanisms, and lower costs. Quality control measures will be used to ensure the collection of high quality data. Enumerators will record respondent information, non-response, and reasons for non-response. Additionally, GPS data will be connected with each survey respondent to ensure that interviews are actually carried out at respondents' households.

Data analysis

46. Following data cleaning, the survey company will carry out descriptive data analysis including a before-after analysis, and a causal (with/without) analysis of the project's impact domains. The analysis will test whether changes occurred overall, in each cluster, and by groups with smaller standard deviations as appropriate for specific variables. In order to carry out before-after analysis, paired sample t-tests will be used to see if individuals have made a significant improvement on the above measures. The use of a t-test rather than a simple comparison of means or percentages of households responding a certain way allows testing whether changes are actual changes or whether apparent changes are likely to be the result of measurement error.
47. To carry out with-without (causal) analysis, propensity score matching (PSM) will be done and difference-in-difference calculations of average treatment on the treated (ATT). PSM will be conducted using the socio-demographic data contained in the household passport section of the survey. Thereafter, attempt will be made to match direct and indirect beneficiaries with those who would also be likely to be direct or indirect beneficiaries based on the question block on participation.
48. Average Treatment Effect (ATE) for the treatment group will be calculated, which should provide estimates of project impact (including magnitude and direction) in all the above project impact domains. Standard errors and/or test for other heterogeneous treatment effects will be carried out by focusing analysis on subgroups and using quantile regression models. If results show statistically significant impacts, tests such as Rosenbaum's robustness tests to determine the

⁹ Where the predominant local language (Swahili) is not spoken, local entrepreneurs who speak the language will be deployed.

sensitivity of results to hidden confounders (i.e., by checking the sensitivity of results to the identifying assumptions underlying the matching analysis) will be conducted.

Qualitative method

49. Although the quasi experimental method suggested is grounded in quantitative approach to evaluation, incorporating qualitative methods will enrich the quality of the evaluation results. Importantly, the qualitative component will provide information and analysis on topics and evaluation criteria for which the quantitative analysis is not suitable and will help probe into issues that emerge from a detailed review of existing project documentation including the reasons for impact – negative or positive, intended or unintended. Where possible, it will also help to assess the possible recall bias generated in the quantitative method. Whilst the household survey will target the farmers, the focus groups will consist of traders, input suppliers and transporters. Adequate representation of women in these groups will be achieved, where possible.
50. The qualitative part of the survey will also be key to identify confounding factors at play which are challenging to control with an ex-post survey data collection. Data collection will take the form of a combination of techniques: focus group discussions (including with men and women separately); individual interactions (interviews with community leaders, key informant interviews with project and government officials) and other techniques such as analysis of documentation, if feasible). A total of 14 Key Informant Interviews (two in each county) comprising traders, transporters and stockists) and 7 Focus Group Discussions (comprising farmers groups – male and female separately) will be undertaken.

Innovative methods to be explored

51. This evaluation will explore innovative methods, relative to the past impact evaluations conducted by IOE. The **Instrumental Variable (IV)** approach is one such method. In order to ensure a reasonable level of confidence in our impact estimates, more than one method will be explored. One such method is the use of Instrumental Variables (IV). The IV approach when used in combination with PSM or DD, can be a good approach for dealing with selection bias due to unobservables. An instrumental variable is a variable that affects an individual's participation in an intervention but does not directly affect outcomes (Deaton, 2009). The IV factor will be used to simulate which participants would have been in the treatment group and which would have been in the control group had the project been based on that factor. The difference in outcomes between these simulated treatment and control groups will constitute the project's impact.
52. The team will explore the use of IV in the Kenya impact evaluation. It will explore the use of the commonly used approach of the Intention-To-Treat (ITT) analysis.¹⁰ A caveat is in order though. For an IV estimation approach to be viable, the IV used must be strong predictor of whether or not a participant will receive the treatment at the same time ensuring that the variables themselves will not determine the Programme's outcome. It will likely be difficult to identify variables that meet both the criteria of IV – a variable that determines participation but does not influence outcomes. The IV approach, however, can lead to worse biases compared to an approach like Ordinary Least Squares (OLS), especially when the identified instruments are not truly exogenous.

¹⁰ Traditional formal statistical analyses focus on the relationship between the assignment and the outcome of interest, discarding entirely any information about the treatment actually received. In other words, the situation where some units who are assigned to receive a particular treatment level do not comply with their assignment. This is common in settings with individual people as the units of analysis, where receipt of the active treatment requires individuals to take a particular action, such as entering a training Programme, as was required under the project.

53. **Bootstrapping.** Although sample size was calculated using power analysis, the sample may still face the problem of lack of power if the effect is disaggregated by commodity chain. To increase the power, bootstrapping will be used. The bootstrap is a general approach to statistical inference based on building a sampling distribution for a statistic by resampling from the data at hand. There are several forms of the bootstrap, and, additionally, several other resampling methods that are related to it, such as jack knifing, cross-validation, randomization tests, and permutation tests.

V. Organization and responsibilities

54. In line with IFAD's evaluation policy, IOE will ultimately be responsible for designing and conducting the impact survey, and for preparing the final evaluation report. The impact evaluation team will be composed of:
- Hansdeep Khaira, IOE Evaluation Officer, who is the lead evaluator for this impact evaluation. He will work under the supervision of IOE Lead Evaluation Officer, Johanna Pennarz. He will be provided technical support by Shijie Yang, IOE Evaluation Research Analyst, and an international consultant, Mr. Matteo Borzoni, with solid quantitative and evaluation skills and good knowledge of commodity value chains, will also be part of the team.
 - The lead evaluator, with inputs from the team, will design the sampling strategy, prepare the survey instruments and ensure the day to day implementation and quality of the impact evaluation. In addition, along with the international consultant, the lead evaluator will undertake the main mission to the country for collecting evidence and information from project stakeholders. These findings, together with the desk review, will serve as the basis for drafting the impact evaluation report.

A local company selected through competitive process, will be tasked with undertaking the data collection part of the impact assessment including translating the survey instruments in local languages, training the enumerators, undertaking the survey, ensuring quality control in the field, compiling data in electronic form; and perform statistical analysis of the data.

VI. Timeline

55. During the preparatory phase, a **first reconnaissance mission** was fielded to Nairobi in September 2017 in order to make contact with the PMU, and identify national organizations with experience in managing quantitative surveys.
56. The **country work** phase will involve the fielding of the impact survey. After the completion of data collection and quality assurance, the econometric analysis and the qualitative analysis will commence. IOE will comment on the preliminary results of the analysis, which will be revised and refined. Thereafter, IOE will field a validation mission and discuss its preliminary results within IFAD and with the programme management and government authorities.
57. During the **report preparation** phase, IOE will draft the main evaluation report, which will be peer-reviewed within IOE and later shared with the IFAD reference group and the external reviewers, as well as with the Government of Kenya for its comments. The revised and final report will be discussed with the Evaluation Committee.
58. In addition to the evaluation carried out, additional work of a methodological nature may be carried out, for instance, further econometric model development using non-experimental methods to compare and validate the results of the quasi-experimental method used, and the preparation of a technical paper.

Table 3
Tentative timeline for the evaluation¹¹

<i>Time</i>	<i>Event</i>
July 2017	Selectivity framework
September 2017	Preparatory mission to Kenya
October 2017 (2nd week)	TORs prepared and sent to IFAD Procurement
November/early December 2017	Draft Approach Paper shared with PMD and SKD
December 2017	Methodology developed, including sampling strategy
December – March 2018	Field survey designed, conducted (pilot and final), data analysed and draft report on impact evaluation prepared
April 2018 (end)	IOE peer review of draft evaluation report
May 2018 (end)	Draft evaluation report shared with IFAD and government of Kenya
June 2018 (end)	Final report available
25 October 2018	Evaluation report presented to the Evaluation Committee
December 2018	Presentation of findings to IFAD and through an in-country event

¹¹ As on 25 January 2018, the activities have been conducted as per the timelines mentioned in Table 3 above.

References

- Baker, Judy L. (2000): *Evaluating the Impact of Development Projects on Poverty, A Handbook for Practitioners*, The World Bank, Washington DC.
- Deaton, Angus S. (2009); *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*; NBER Working Paper No. 14690, January 2009.
- Hopkins, Daniel and King, Gary (2010): *Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability*, Public Opinion Quarterly, 1-22.
- IEG (2011): *Impact Evaluations in Agriculture, An Assessment of the Evidence*, IEG, The World Bank, Washington DC
- IFAD (2015): *Evaluation Manual, Second Edition*, Independent Office of Evaluation of IFAD.
- ILRI (2014): *Evaluating value chain interventions: A review of recent evidence*, International Livestock Research Institute (ILRI) Discussion Paper 26.
- Khandker, Shahidur R.; Koolwal, Gayatri B.; Samad, Hussain A. (2010): *Handbook on Impact Evaluation, Quantitative Methods and Practices*, The World Bank, Washington DC.
- King, Gary; and Wand, Jonathan (2007): *Comparing Incomparable Survey Responses: New Tools for Anchoring Vignettes*, Political Analysis 15: 46–66
- Michael Bamberger (2000): *Integrating Quantitative and Qualitative Methods in Development Research*, World Bank and Weiss op. cit.)
- Rosenbaum, P. and Rubin, D. (1983): *The Central Role of the Propensity Score in Observational Studies for Causal Effects*. Biometrika, 70, pp. 41-50.
- The Millennium Challenge Corporation (2012): *Principles into Practice, Impact Evaluations of Agriculture Projects*.
- Winters, Paul; Salazar, Lina; Maffioli, Alessandro (2010): *Designing Impact Evaluations for Agricultural Projects* Impact-Evaluation Guidelines, Technical Notes, No. IDB-TN-198, Inter-American Development Bank.

Selectivity framework for the Impact Evaluation

The projects to be selected for impact evaluations by IOE are guided by a comprehensive selectivity framework. The purpose of the selectivity framework is to enhance the transparency in the selection and prioritization of projects for IOE impact evaluations.¹²

Essential criteria

Criteria	Guiding questions for IEs	Answers
<i>Evaluation results for learning</i>	Is this a country where IOE will conduct a CPE in 2018/2019?	YES
	Will the findings of this IE, given the sub-section nature of the project, also feed into on-going or planned evaluation synthesis reports or Corporate Level Evaluations (CLE) by IOE?	YES (CLE on Value Chain)
<i>Project status</i>	Did the project implementation end between 1 and 3 years ago?	YES
<i>Geographical distribution</i>	Has IOE conducted an interim or completion evaluation or PPA on this project in the past?	NO
	Is this a project where IFAD (RIA) is undertaking an impact evaluation by the end of 2018?	NO

Technical criteria

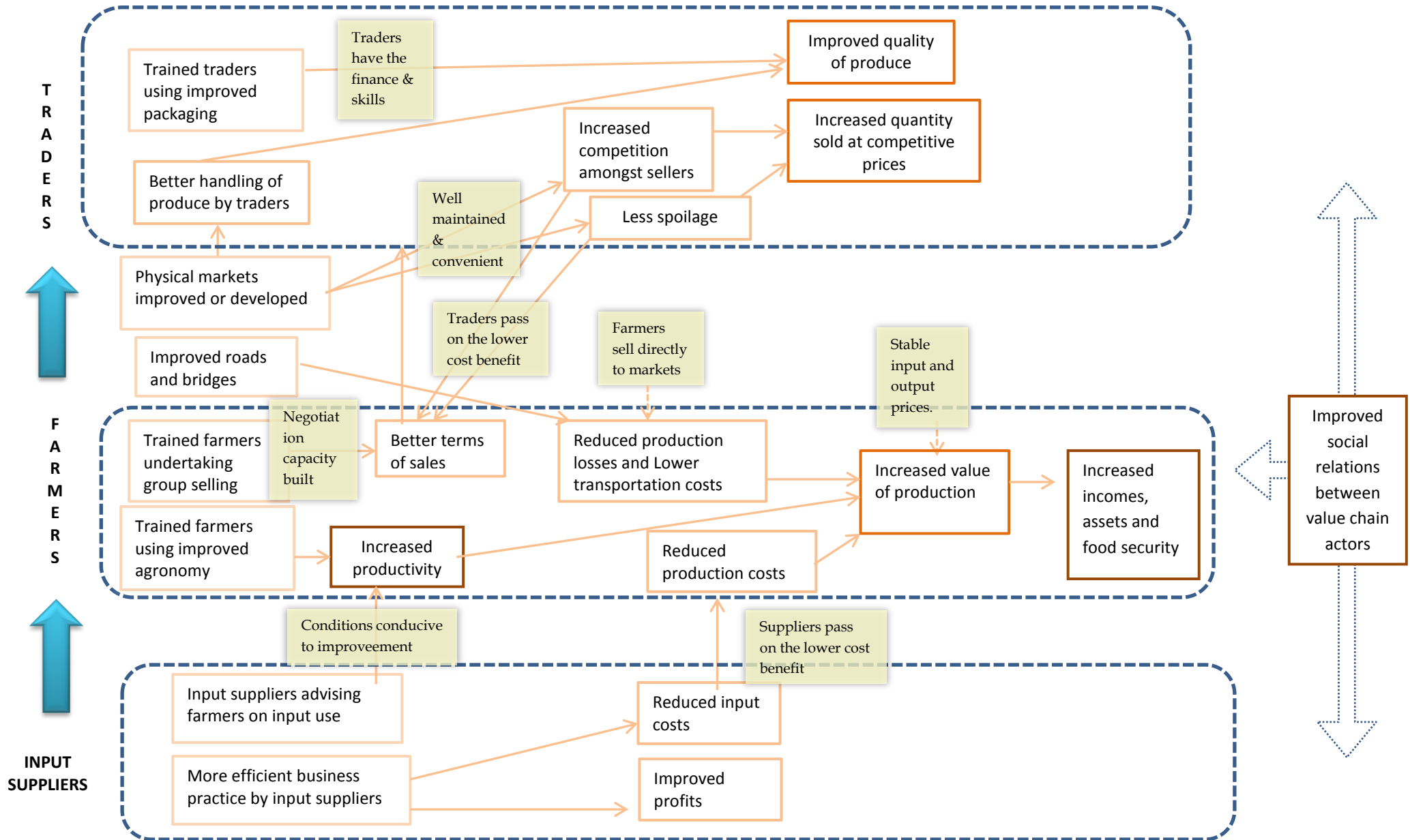
	Is a baseline survey available? What is its quality? Did it include control or comparison groups? Is an electronic database available?	Yes Satisfactory No No
	Is a RIMS baseline survey available? ¹³ What is its quality? Did it include control or comparison groups? Is an electronic database available?	No N/A N/A N/A
	Is a RIMS completion survey available? What is its quality? Did it include control or comparison groups? Is an electronic database available?	NO N/A N/A N/A
	Are other surveys available? If so: What is their quality? Did they include control or comparison groups? Is an electronic database available?	Completion Not satisfactory Yes Yes
	What is the quality of the PCR including in terms of data and analysis on impact?	To be determined

¹² Based largely on the selectivity framework, IOE will normally undertake impact evaluations of projects: (i) within three years of their completion date; (ii) that are not selected for impact evaluation by IFAD Management; (iii) that will also be included as part of the project portfolio analysis in forthcoming CSPes, to enhance the latter's evidence base; (iv) that have innovative development approaches (e.g. institutional, social, technological) that merit deeper analysis and documentation; and (v) that offer enhanced opportunities for learning, on what works and what does not in promoting sustainable and inclusive rural transformation.

¹³ The baseline and endline surveys are not RIMS compliant in that there is no asset index (poverty) and no child malnutrition (anthropometric measures) or length of hungry season (food security), although poverty is assessed through income and food security through consumption of different food categories.

<i>Evaluability assessment</i>	Is a MTR available?	Yes
	What is the quantity and quality of data generated by the project's M&E system?	Reasonably good
	What is the availability and quality of project logical framework in President's Report?	Reasonably good quality
	Are qualitative thematic studies available?	Yes, some
	Did the project experience implementation delays?	No
<i>Availability of local technical expertise</i>	Is national technical expertise in quantitative and qualitative data collection and analysis available?	Yes

Theory of Change (value chain)



Logical Framework

Results Hierarchy	Indicators a/	Means of Verification
Goal: Contribute to reduced Poverty and improved health among poor rural households in medium-high potential horticultural farming areas	<ul style="list-style-type: none"> • 10% reduction of poverty prevalence rate among 12,000 households participating in the project by Year 7 (Baseline 35% in 2003)RIMS3 • 3% reduction in malnutrition prevalence (weight for age of children under 5) in project area by Year 7 (reduction in chronic malnutrition – 36% in 2003, underweight 17% in 2003 and wasting 6% in 2003) • 5% increase in inventory of household assets among 12,000 participating households in project area by Year 7 (Baseline 35% in 2003) 	<ul style="list-style-type: none"> • Household income and expenditure surveys. • RIMS impact survey questionnaire (baseline and final) • Demographic and health surveys conducted by Kenya National Bureau of Statistics • Annual household asset surveys by M&E
Development Objectives: Increased domestic horticulture productivity and improved functional input and produce marketing system	<ul style="list-style-type: none"> • 10% increase in average real incomes for 12,000 households engaged in Sustainable domestic horticulture enterprises by Year 7 (Baseline to be determined). • 10% increase in value of marketed horticultural produce by year 7 (Baseline). • 10 % Increase yield per ha (Baseline to be determined). • 10 %Net margin per unit area (Baseline to be determined). • 5% Increase in unit price for producers (<i>Baseline for unit price for producers to be determined</i>). • 10% decrease in price of inputs (suppliers and producers) (Baseline for input prices to be determined). 	<ul style="list-style-type: none"> • Baseline and annual production and income surveys in Project area by M&E and during impact survey in Year 5. • Specific evaluation studies
Outcome A1: Informed Investment Decision	<ul style="list-style-type: none"> • Number of community projects implemented (RIMS2) ,(Baseline 0 in 2007) target 80 in year 7 implemented (RIMS2) 	<ul style="list-style-type: none"> • Annual Project reports. • infrastructure registers • PMU Assessment
Output A1.1 Analytical Studies conducted	<ul style="list-style-type: none"> • 14 No. VCA studies conducted ,(Baseline 0 in 2007) target 14 in year 7 • Upstream/downstream/Price stability study conducted,(Baseline 0 in 2007) target 1 in year 7 • Number of community action plans included in local government plans ,(Baseline 0 in 2007) target 80 in year 7 implemented (RIMS2) • Number of community action plans formulated and implemented(RIMS1), target of 80 in year 7, Baseline 0 in 2008 	<ul style="list-style-type: none"> • Contract register • study reports
Outcome B 1: Empowered Horticulture System Actors	<ul style="list-style-type: none"> • 10 % increase of market actors benefiting from improved market access. (Baseline to be determined). • 10 % increase Volume of business per unit enterprise (Baseline to be determined). 	<ul style="list-style-type: none"> - Impact assessment survey report - Baseline survey reports
Output B1.1 Capacity of GoK Staff in marketing systems improved	<ul style="list-style-type: none"> • Government officials trained (RIMS) Baseline of 0 in 2007 (target of 2000 by Year 7). 	<ul style="list-style-type: none"> • DAO Progress report • Infrastructure register • group register
Output B1.2 Capacity of value chain players in marketing service provision Improved	<ul style="list-style-type: none"> • No. of value chain players trained by category:- • People trained in post production, processing and marketing (RIMS1) (target of 12,000) • Marketing groups formed and/or strengthened (RIMS1) (target of 600 by Year 7) • Marketing groups with women in leadership positions (RIMS1)(target of 200 by Year 7) 	<ul style="list-style-type: none"> • DAO Progress report • Infrastructure register • group register

Results Hierarchy	Indicators a/	Means of Verification
Output B1.3 Existing Marketing Information Systems improved	<ul style="list-style-type: none"> Percentage of value chain players accessing timely and reliable market information (60% compared to a baseline of 12% in year 2007) 	individual enterprise report
Output B1.4 Access to affordable financial services supported	<ul style="list-style-type: none"> Enterprises accessing facilitated financial services (RIMS 1) (baseline of 0, target of 5,000 by year 7) Value of gross loan portfolio (RIMS 1) (target of KES 1 billion by Year 7) 	<ul style="list-style-type: none"> District Reports Equity Bank
Output B1.5 Development of legal and regulatory environment for input and produce (policy Developed) facilitated	<ul style="list-style-type: none"> National Horticulture Policy developed(baseline of 0 in 2007, target of 1 by year 7) Number of pro-poor legislation and regulations enforced at the local or central level (RIMS 2) ,(baseline of 0 in 2012, target 2 of by year 7) 	<ul style="list-style-type: none"> ASCU report Sub-County reports
Outcome C1. Developed sustainable marketing Support Systems	<ul style="list-style-type: none"> Number of functioning infrastructure (RIMS2) (target 60 by Year 7). 	<ul style="list-style-type: none"> DAO Progress report Infrastructure register group register
Output C1.1 Innovations in value addition and market oriented production technologies enhanced	<ul style="list-style-type: none"> Number of pilot initiatives supported by category (Baseline of 0 in 2007, target of 80 by year 7) Numbers of innovations adopted/replicated (RIMS2) (Baseline of 0 in 2007) 	<ul style="list-style-type: none"> DAO Progress report Infrastructure register group register
Output C1.2 Rural access roads improved	<ul style="list-style-type: none"> Number of roads improved (target of 92 by Year 7). Baseline 0 in 2008) Length of rural roads opened up through spot repairs (target of 230 km by year 7) baseline of 0 in 2007 	<ul style="list-style-type: none"> Infrastructure register Baseline survey
Output C1.3 Physical market infrastructure improved	<ul style="list-style-type: none"> Number of market facilities developed/improved(RIMS) (target of 50 by year 7, baseline of 0 in 2008) Volumes of priority crops traded ,(baseline of --in 2007, target of -- by year 7) Environmental management plan formulated (RIMS 1)(baseline of 0 in 2008, target of 72 by year 7) 	<ul style="list-style-type: none"> EIA report Infrastructure register Local authority records Infrastructure register
Outcome D 1: Effective and efficiently managed Programme	<ul style="list-style-type: none"> Project activities fully integrated in mainstream GoK systems and institutions with functional management, monitoring and reporting (target of --- by Year 7) 	<ul style="list-style-type: none"> NIMES M&E reports
Output D 1.1: Fully functional governance, management, monitoring and reporting systems.	<ul style="list-style-type: none"> Project implemented on schedule with performance ratings of satisfactory or better. Increasing measures of institutional capacity. 	<ul style="list-style-type: none"> Supervision and implementation support mission reports, and audit reports. Formal institutional capacity assessments
Output D 1.2: Knowledge about NRM effectively managed and disseminated to stakeholders.	<ul style="list-style-type: none"> Increasing dissemination and use by stakeholders of knowledge generated by Project. Regional knowledge centres effectively networked. 	<ul style="list-style-type: none"> Number of information materials produced and distributed project-wide as monitored by M&E. Reports of regional knowledge networks. Surveys on awareness of sustainable NRM.

Evaluation Framework

Evaluation criteria	Key evaluation questions	Performance Indicators	Data Sources
Rural poverty impact	Was there an improvement in the socio-economic situation of beneficiaries? Were beneficiaries below poverty line lifted out of poverty? What elements of the project were most important in creating the desired outcomes? What other impacts (positive or negative) did the intervention have on the wider community?	USD1.25/day or national poverty line	Household survey Focus Group Discussion Secondary data
<i>a) household income and assets</i>	What have been the changes in incomes and assets in the beneficiary group as compared to the non-beneficiaries and with respect to project baseline? Are the productive investments increasing in project areas? What factors caused the above changes?	Wealth quintile Household expenditure (food, non-food) Increase in the number of sources of income	Household survey Focus Group Discussion
<i>b) human and social capital and empowerment</i>	To what extent the project contributed to strengthening the role of community based organizations in development activities? To what extent did the behaviour of the communities change towards the adoption of sustainable agricultural practices?	Participation in village groups and associations Women's participation in village groups and associations	Household survey Focus Group Discussion
<i>c) food security and productivity</i>	What have been the changes in the food security situation, including nutrition, of beneficiaries? Did farmers adopt new agricultural practices as a result of the project? Did the project interventions affect agricultural and livestock productivity?	Increased intake/dietary diversity. Average area under production, average yields per hectare.	Household survey Focus Group Discussion
<i>d) institutions and policies</i>	What are the changes in the quality and performance of institutions, policies and the regulatory framework that influence the lives of the poor? What has been the project's contribution to the behavioural changes in local authorities and grass roots organizations? What were the underlying causes for the induced changes?	Influence on policies/practices concerning leasing.	Household survey Focus Group Discussion
<i>e) natural resources, and environment</i>	What has been the impact on natural resources and environment? To what extent and how did the project contribute to the sustainable use of water?	Soil and water management, sustained production under climate variability.	Household survey Focus Group Discussion
Project performance			
Relevance	Were the objectives of the project relevant to: i) country strategies and policies? ii) the needs of the beneficiaries? iii) IFAD's priorities, strategies and COSOPs? Was the project design based on a thorough socio-economic analysis of the sector, including gender related aspects? Did it target the poorest communities, including women? Was it based on development approaches tailored to the context? Did the project have an exit strategy at design?	Proxy indicators of relevance	Desk review Key Informant Interviews
Effectiveness	Was the project targeting approach effective?		Desk review

	<p>What was the project outreach at completion? Did the project meet its objectives? For instance, did farmers adopt new agricultural practices as a result of the training? Did rehabilitating roads and bridges increase access to markets? Did the above result in increased crop production and/or crop diversification? Did they sell more due to coming together as groups?</p>		<p>Household survey Focus Group Discussion Secondary data Key Informant Interviews</p>
Efficiency	<p>How economically resources and inputs were converted into results? Were the project effects large enough to justify its costs? (Economic Rate of Return)? What was the time lag between approval and loan effectiveness? What was the budget utilization at completion? Were the funds from IFAD and other partners made available in a timely manner? What are the project management costs at completion? And compared to other similar projects?</p>		<p>Desk review Key Informant Interviews</p>
Sustainability of benefits	<p>To what extent are net benefits deriving from the project continuing? To what extent did the project contribute to reduce the vulnerability of the sector? What is the sustainability of the project from a technical, institutional and social perspective? Is there evidence that the infrastructure investments will be sustained after rehabilitation was completed?</p>		<p>Household survey Focus Group Discussion</p>
Other criteria			
Gender equality and women empowerment	<p>Did the project expand women's access to and control over fundamental assets? Did the project strengthen women's agencies – their decision-making role in community affairs and representation in local institutions? Did the project improve women's well-being and ease their workloads by facilitating access to basic rural services and infrastructures?</p>		<p>Household survey Focus Group Discussion</p>
Innovation	<p>To what extent did the project introduce innovative approaches?</p>		<p>Desk review Key Informant Interviews</p>
Scaling-up	<p>Have any approached of the project been scaled-up by the government and others? To what extent did the project learn from past experience and inform the design of new projects?</p>		<p>Desk review Key Informant Interviews</p>