

## **Approach Paper**

# **Impact Evaluation of the Agricultural Support Project in Georgia**

January 2017

## Contents

### Abbreviations and Acronyms

1. Background .....	3
2. Project Description.....	4
3. Preliminary evaluability assessment of the project .....	6
4. Methodology .....	7
<i>The evaluation approach .....</i>	<i>7</i>
<i>Intervention logic of the project.....</i>	<i>8</i>
<i>Mixed methods approach .....</i>	<i>8</i>
<i>Quantitative research design .....</i>	<i>9</i>
<i>Qualitative method.....</i>	<i>14</i>
5. Core Learning Partnership.....	15
6. Organization and responsibilities.....	15
7. Timeline .....	16
Annex 1 References.....	18
Annex 2 Selectivity framework for the Impact Evaluation .....	19
Annex 3 Reconstructed Project Intervention Logic.....	21
Annex 4 Logical Framework.....	22
Annex 5: Evaluation Framework.....	24

### Abbreviations and acronyms

IOE	Independent Office of Evaluation of IFAD
ASP	Agriculture Support Project
IE	Impact Evaluation
M&E	Monitoring and Evaluation
PFI	Participating Financial Institutions
TA	Technical assistance
RIMS	Results and Impact Management System
PSM	Propensity Score Matching

# Impact Evaluation of the Agricultural Support Project in Georgia

## Approach Paper

### 1. 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. In addition to contributing to the repository of impact evaluations, each successive IE harnesses internal learning by taking cognizance of the experience of its predecessor in its design<sup>1</sup>. In 2016-2017 the office is undertaking its fourth impact evaluation. The programme selected for the impact evaluation is the Agricultural Support Project (ASP) in Georgia. The project was selected using a comprehensive selectivity framework (Annex 2).

2. This Approach Paper lays out the terms of reference for the ASP 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 for the design of future projects. The following are its main objectives:

- i) To measure, and in doing so, establish if the project interventions had a welfare 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 assess the innovative features of the project's design and provide the information needed to scale up successful project components and to inform the design of similar projects in future, thereby strengthening project effectiveness.
- iii) To provide useful evidence for and to be used as a critical input towards the Georgia country programme evaluation planned by IOE in 2017.

3. The results of the evaluation are expected to contribute to better informed decision-making and learning about successful approaches to increase incomes and reduce poverty and to promote greater accountability for performance of IFAD's projects. This IE will add to IFAD's database of impact evaluations including those carried out by its Strategy and Knowledge Department in collaboration with the Programme Management Department. To this end, it 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. Although the principal aim of this evaluation is to assess the impact of the project on welfare, the evaluation will also assess key evaluation criteria

---

<sup>1</sup> This impact evaluation builds on IOE's previous experience with impact evaluations in Sri Lanka, India and Mozambique.

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. Other areas to be assessed will include: 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).

## 2. Project Description

5. The Agricultural Support Project (ASP) was designed as a four-year project to be implemented between 2010 and 2014. The project was extended by one year to compensate for start-up delays and to ensure completion of vital works and formally closed in December 2015<sup>2</sup>. Below is a description of the project.

Key Dates					
IFAD Approval	Signing	Effectiveness	Mid-Term Review	Original Completion	Actual Completion
17 <sup>th</sup> December 2009	8 <sup>th</sup> July 2010	8 <sup>th</sup> July, 2010	7 <sup>th</sup> June 2013	30 <sup>th</sup> September 2014	30 <sup>th</sup> September 2015
Mid-term Review	Interim Evaluation	Original Loan Closing	Actual Loan Closing		
7 <sup>th</sup> June 2013		31 <sup>st</sup> March 2015	31 <sup>st</sup> December 2015		
IFAD Financing					
Loan	USD	13,500,000	% disbursed	76%	
Grant	USD	200,000	% disbursed	97%	
Actual Costs and Financing (USD '000)					
Component	IFAD	PFI's	Beneficiaries	GOVT	Total
Rural Leasing	1,809,173	558,048	461,631		2,621,363
Rural Infrastructure	7,944,819		745,634	1,238,655	9,747,973
Project Management	661, 267				700,644
Total	10,415,259	558,048*	1,207,265	1,447,416	13,069, 940
<b>Remarks:</b> The co-financing envisaged for the project did not materialise, and IFAD approved a supplementary loan of USD 5 million to make up the shortfall.					

**Target Population:** The project was aimed at agriculture-related producers and processors among poor rural women and men.

**Goal:** The overall goal of the project was to increase income among poor, small-scale but increasingly commercially oriented farmers and rural entrepreneurs.

**Objectives:** The project's objectives were: (i) to increase assets and incomes among actually and potentially economically active poor rural women and men willing to move towards commercially viable agricultural and associated rural enterprises; and (ii) to remove infrastructure bottlenecks which inhibit increasing

<sup>2</sup> According to project documentation, in practical terms, project activities commenced only in 2012.

participation of economically active rural poor in enhanced commercialization of the rural economy.

*Components:* The project had two main components, 'Support for Rural Leasing' and 'Small Scale Rural Infrastructure' which included eight of the nine regions in Georgia and covered 13 of the 83 municipalities in the country. Except in the case of two municipalities, there was no overlap of geographic coverage i.e. both components were employed in separate municipalities.

#### *Component 1: Rural Leasing*

The Rural Leasing component included two sub-components; (i) Capital to refinance leasing contracts of Participating Financial Institutions (PFI) and (ii) International technical assistance (TA), trainings and exchange visits. The component aimed at recapitalisation through financial leasing and consequent modernisation of Georgian agriculture, specifically among poor smallholders, small and medium agro-related enterprises. The use of agro-leasing and rural micro-leasing was an innovative feature for both IFAD and Georgia and was expected to channel additional financial resources to the agriculture sector and to help reduce rural poverty.

#### *Component 2: Small Scale Rural Infrastructure*

Through the component, ASP was expected to award competitive contributory grants for investments in public infrastructure that would enable and enhance the rural population's investments and activities in on-farm and off-farm related business. The schemes consisted of the rehabilitation of three rural infrastructure schemes including two deteriorated bridges designed to facilitate transport and communication of agricultural products and the movement of livestock to the summer pastures and one drinking water supply system to make better use of available water resources from four springs. The irrigation schemes include the rehabilitation of six schemes covering an area of about 11,042 hectares.

6. The project was expected to lead to greater incomes among poor, small-scale but increasingly commercially oriented farmers and rural entrepreneurs. Benefits would derive from: (i) crop intensification and increased production of high value products through better irrigation; (ii) improved small-scale rural infrastructure for the benefit of those engaged in agriculturally related activities; (iii) reduced transportation costs and post-harvest losses due to rural road rehabilitation; (iv) increased productivity and reduction of crop wastage and production costs through improved mechanisation and availability of modern equipment due to provision of leasing contracts and reduced collateral constraints; and (v) incremental tax revenues as a result of increased volume of taxable production.

7. Targeting: The ASP had an element of geographic targeting in that, although the project had national coverage, the priority was given to areas with the highest concentration of poor rural people. In addition, self-targeting was manifested in: i) sector specificity – agriculture or agro-related; ii) scale – upper limits for investments making them unlikely to be of interest to richer socio-economic categories and appealing to relatively poor rural women and men; and, iii) modality i.e. the requirement of backward linkages in terms of value chains of importance to poor primary producers and employment generation

among poor rural people. Thirdly, the project used direct targeting through channelling resources to individuals, households or groups identified according to specific eligibility criteria. These criteria included: i) preference to those owning/operating less than one hectare of land; ii) preference to those registered for Targeted Social Assistance; and, iii) minimum 30 per cent representation of women in all categories of project investments<sup>3</sup>.

8. Table 1 shows details of project outreach as per the two project components.

Table 1: Beneficiaries by project component (end of project)

<b>Rural Infrastructure</b>		<b>Rural Leasing</b>	
Irrigation	14453	Enterprises	15
Drinking water	500	Enterprise owners	41
Bridges	540	Employment	1152
Labour employed during construction	300	From backward linkages created (indirect beneficiaries)	2645

Note: data refer to households (except in the case of enterprises)

### **3. Preliminary evaluability assessment of the project**

9. 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. This section provides results of the preliminary assessment for ASP.

10. The project had undertaken both baseline and endline studies. However, since the Monitoring and Evaluation system was established two years after the project commencement, the baseline was undertaken after project start. The baseline study was conducted in the project treatment area only (sample size of 900 households) and used two-stage stratified cluster sampling. The end-line was conducted using quasi-experimental method with a control/comparison group and using the difference-in-difference method (450 households in the treatment area and 450 households in the control area). Similar to the baseline survey, the end-line survey employed household survey as the data collection tool. Sampling weights were used to ensure representativeness.

11. The project Monitoring and Evaluation unit has the list of all villages where the project was implemented; lists of project beneficiaries is not available since all households in the villages were considered to be targeted (assuming that 88 percent of all agriculture holdings in Georgia are less than 1 hectare). List of all enterprises targeted is available (leasing component). RIMS data are available for all years and for levels 1 (outputs) and 2 (outcome ratings).

<sup>3</sup> Further details on the project can be found using the follow link that also provides results achieved at different stages of the project (through supervision reports including for mid-term review):

<http://operations.ifad.org/web/ifad/operations/country/project/tags/georgia/1507/documents>

12. Based on analysis of available project data, several shortcomings have been observed, such as:

- The sample size was not determined using established and recognised methods i.e. power calculations.
- The comparison group was selected during the endline only, with recall method used to construct the baseline information for this group.
- Techniques for matching observables or unobservable characteristics of respondents were not employed. The comparison group was selected based on neighbouring settlements where no ASP project activities have been implemented.
- Food security was assessed using some basic questions on food deprivation.
- Household survey was used as the only instrument, with no qualitative studies being employed.
- Only direct beneficiaries of the leasing component (agro-enterprises) were surveyed; indirect beneficiaries such as those receiving employment, farmers supplying produce to direct beneficiaries, etc., were not surveyed.
- A high level of non-response for income-related variables.

#### **4. Methodology**

##### *The evaluation approach*

13. In light of the results of the 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.

14. As mentioned earlier, in addition to evaluating the four impact domains (under the rural poverty impact criterion) viz.,: (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)<sup>4</sup>.

15. 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.

---

<sup>4</sup> 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>

## *Intervention logic of the project*

16. The intervention logic of the project (or, 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 project intervention logic 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. One more objective of the intervention logic is to present the assumptions that underpin the transition along the causal path and which mainly concern behavioural change on the part of beneficiaries.

17. The figure in the annex shows the causal path for the three main interventions carried out by the project viz., rehabilitation of bridges, rehabilitation of irrigation canals and rural leasing to enterprises. Thus, for example, financing for leasing by rural enterprises results in increased investment by lessees into machinery or equipment (output) which in turn leads to increased demand for agricultural/livestock products (as raw materials) and for workforce (immediate outcomes). As depicted by the figure, the three interventions lead to common medium-long term outcomes such as increased production by the enterprise-sourced farmers and increased and/or more stable employment. The impact or the goal of the project is an increase in the general well-being of beneficiaries driven largely by increase in incomes and consumption (food and non-food).

18. 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.

## *Mixed methods approach*

19. Good evaluations are almost invariably mixed method evaluations<sup>5</sup>. 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. These mixed methods are based on the principle of *method triangulation* i.e. the use of multiple methods to study a single problem or program, such as interviews, observations, questionnaires or written secondary sources. The two methods can be carried out either contemporaneously or sequentially. In the case of this impact evaluation, these will be undertaken in parallel, for reasons of cost and time efficiency.

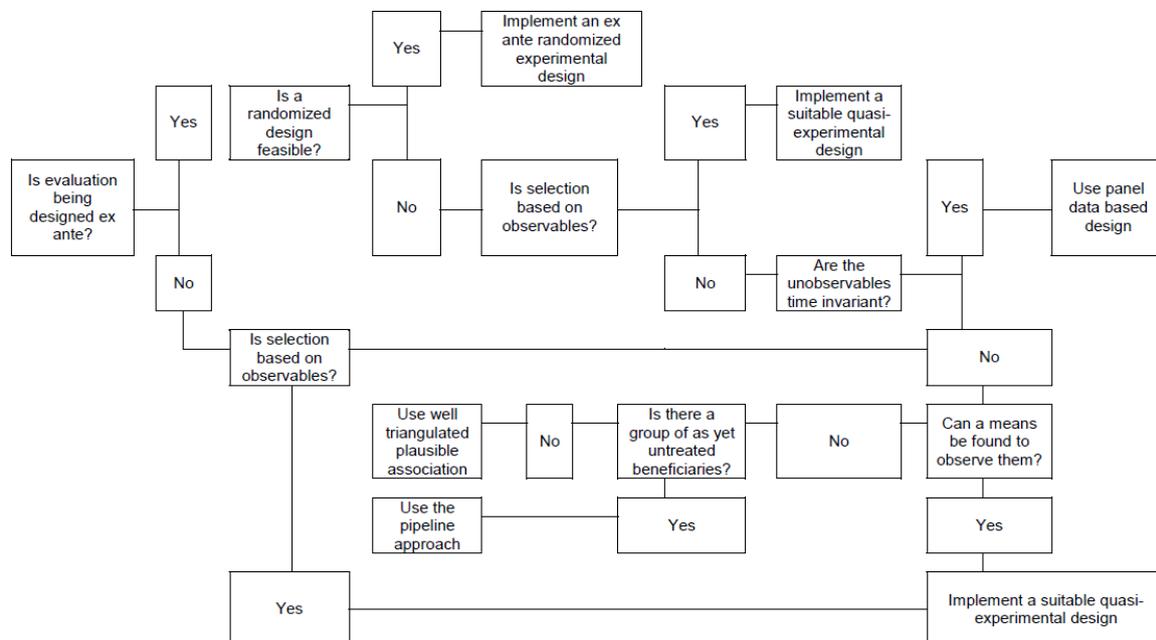
---

<sup>5</sup> Michael Bamberger (2000) Integrating Quantitative and Qualitative Methods in Development Research, World Bank and Weiss op. cit.)

20. Since the study is ex post, a panel is not possible, and since selection into the project could have been determined by unobservables, the problem of selection bias could remain. To overcome this, the evaluation will rely on program theory, as described earlier, to build an argument by plausible association, and relevant quantitative and qualitative methods, are described below. 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.

21. 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 (Figure 1 below shows the decision tree for selection of evaluation design). 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.

**Figure 1: Decision Tree for selecting Evaluation Design**



Source: NONIE (2009)

## Quantitative research design

### Household Survey

22. 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. Additionally, the recall method will be used, collecting data on respondents' past behaviour, to recreate a baseline scenario, thus allowing making approximate before-after comparisons.

23. 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 irrigation sub-component was implemented only in areas with irrigation potential. 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 rural leasing component, in which only some enterprises decided to participate. Those who chose to participate could differ from those who did not.

24. 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, the following methodology will be adopted. First, propensity score matching (PSM) will be undertaken wherein the calculation of propensity scores<sup>6</sup> 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.

25. Second, use of genetic matching (multivariate matching on pre-treatment confounders *plus* the propensity score)<sup>7</sup> to calculate weights for each covariate. Finally, use of multivariate matching to identify which municipalities and clusters are most like the treated ones at the start of project activities.

26. Matching will be done at the municipal level<sup>8</sup> and then at the cluster level. This process will lead to the identification of communities which will be good comparisons to the treated communities. Propensity scores for each municipality will be calculated using the following characteristics from baseline year. Since there are no publicly available data on agriculture or irrigation at the municipal level in Georgia, a number of variables that can serve as proxy to poverty, etc. will be sought from with the National Statistics Office of Georgia.

- Population size
- Ethnic composition
- Religious composition

---

<sup>6</sup> 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.

<sup>7</sup> Genetic matching is a multivariate matching method that uses an evolutionary search algorithm to determine the weight each covariate is given, and improves balance on observable covariates in quasi-experimental studies.

<sup>8</sup> A municipality is a subdivision of Georgia, consisting of a settlement or a group of settlements, which enjoys local self-government.

- Gender composition
- Social assistance package recipients as share of population
- Targeted social assistance subsidy recipients as share of population
- Number of school students as share of population
- Distance from administrative boundary line with the disputed territories of South Ossetia and Abkhazia;
- Predominant settlement type (% rural)
- Monthly need-based subsistence subsidies (per capita, family, registered needy family)
- Number of internally and geographically displaced (per capita)
- Economic composition by business size (ratio of micro to small; small to medium sized businesses)
- Agriculture related data that the municipalities maintain.

27. After identifying matching municipalities, matching at the cluster level will be performed. Electoral precinct boundaries will be used as clusters, due to a lack of territorial boundaries of communities, and predominantly urban clusters will be removed from the sampling frame<sup>9</sup>. Using the same three step procedure described above, matching will be performed at the cluster level. In addition, use will be made of the Köppen Climate Classification System in the identification of clusters<sup>10</sup>.

28. Some of the potential characteristics to be matched at the cluster level will be:

1. Median altitude
2. Distance from urban centre
3. Distance from a significant road
4. Size of irrigation scheme (500-2500 ha of command area)
5. Areas with irrigation potential

29. 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.

30. 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. In addition, several different starting points will be used as opposed to a single one.

---

<sup>9</sup> To sample in Georgia (due to the unavailability of a sampling frame that lists individuals) it is necessary to use a list of territorial units or clusters, as a sampling frame. In Georgia, because the National Statistics Office does not provide the territorial boundaries of communities used for the census, the only practicable option for selecting clusters is to use the boundaries of electoral precincts, which are publicly available. If within the framework of the project, the National Statistics Office (Geostat) would be willing to provide us with the boundaries of the communities, then this list of clusters will be used instead of electoral precincts, because they are likely to more accurately match up with the boundaries of the treatment communities. However, experience dictates that Geostat cannot provide this information to the public due to data privacy legislation.

<sup>10</sup> The Köppen Climate Classification System is the most widely used system for classifying the world's climates. Its categories are based on the annual and monthly averages of temperature and precipitation.

31. The survey instrument will identify whether the individual is a direct or indirect beneficiary. 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.

*Power calculations and sample size*

32. 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. The formula is as follows:

$$n_B = \left(1 + \frac{1}{k}\right) \left(\sigma \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\mu_A - \mu_B}\right)^2$$

33. Assuming standard error levels of alpha=0.05, a beta equal to 0.2, and equal sized treatment and control groups, and from the baseline data deriving a 2 per cent effect size and standard deviation (GEL 200) and mean of household income (GEL 1142), a sample size of approximately 3000 households will be obtained, with 1250 in the treated group and 1750 in the control group. The oversampling of the control group is in order to find the highest quality matches possible for the treated group. However, attempt will be made to explore scenarios for the sample size calculation assuming different effect sizes.

*Survey instrument*

34. The survey instrument will contain question blocks that will enable to assess project impact on household income and net assets, human and social capital and empowerment, food security and agricultural productivity, and environment and natural resource management.

35. Given that the baseline data is insufficient, the recall method will be used in order to reconstruct the baseline. 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. Although the project was officially launched in 2010, project documentation states that project activities started in earnest in 2012, and hence 2012 could be a more appropriate baseline year. In order to attempt to increase the quality of responses, use will be made of vignettes and initial questions requiring the respondent to recall the past in order to stimulate respondent memory, in line with the larger literature on vignettes which are seen as valuable in priming and focusing respondent thinking (King, Gary and Wand, Jonathan 2007).

36. The project had a number of different types of beneficiaries. The survey instrument will identify the beneficiary type in treated communities using a question block that identifies whether they benefited from specific components of the project, directly or indirectly. In the control communities, comparable questions will be asked, but rather than ask about actual participation/inclusion in the project's intervention, the questions will ask about whether the individuals would be interested in participation in such an intervention and how they might participate if a comparable project was scaled to include their community. This set of questions will enable more accurate matching during the analysis phase as well as be useful in informing IFAD's future operations in Georgia about potential scaling of project components. In addition to the aforementioned 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 contagion of interventions.

#### *Fieldwork*

37. Pilot interviews. Once the questionnaire has been developed, some 45-50 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.

38. 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.

39. 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 cleaning will be carried out and probability weights calculated and added to the dataset following fieldwork. The results of this process will be recorded in the technical report.

#### *Data analysis*

40. 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.

41. To carry out with-without (causal) analysis, propensity score matching (PSM) will be done, together with genetic matching and difference-in-difference calculations of average treatment on the treated (ATT) and in the case of gender and other group-related indicators localized average treatment effect (LATE). 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.

42. Following the calculation of propensity scores, this information as well as the propensity scores will be inputted into a genetic matching algorithm. The algorithm will reduce bias on observable covariates. Then, 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 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*

43. Although the quasi experimental method suggested is grounded in quantitative approach to evaluation, incorporating qualitative methods will enrich the quality of the evaluation results and aid in the deeper interpretation of results obtained from a quantitative approach by shedding light on the processes and causal relationships. The qualitative component will provide information and analysis on topics 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. It will also help to assess the possible recall bias generated in the quantitative method.

44. 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 participatory 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).

#### *Other methods*

45. In addition to the above methods, attempt will be made to analyse the impact of the project through the use of geo-referenced data. For instance, the Normalized Difference Vegetation Index (NDVI), that compares average

vegetation in an area with the vegetation in a particular point in time (month, for example), can help assess the change in vegetation cover in project areas due to improved/increased irrigation. Similarly, change in vegetation cover in hitherto areas not frequented by livestock will also help assess the impact of bridges that were constructed for facilitating movement of livestock for grazing to those areas. In conjunction with "ground-truthing" through the household survey, use of NDVI will help in validating results. Depending on data availability and cost of data acquisition, and availability of suitable expert, a final decision on use of geo-referenced data will be taken.

## **5. Core Learning Partnership**

46. Stakeholders' participation is crucial for successfully conducting evaluations in general, and impact evaluations, in particular. This will ensure that the key concerns of the stakeholders are taken into account, that the evaluators fully understand the context in which the project was implemented, and that opportunities and constraints faced by the implementing institutions are identified.

47. In accordance with the Evaluation Policy, a core-learning partnership (CLP) will be established to enhance the quality of the impact evaluation as well as to build ownership among key partners in the evaluation process and its outcomes. The CLP will comprise the following members:

### **i. Representatives of IFAD management**

- The Country Programme Manager for Georgia
- Programme Management Department, Front Office
- Strategic and Impact Assessment Division

### **ii. Government authorities at national level**

- Director and Deputy Director of Ministry of Agriculture
- Director, Agricultural Development Projects Coordination Centre (ADPCC)

### **iii. IOE**

- Deputy Director, F. Felloni
- Lead Evaluation Officer, M. Torralba

## **6. Organization and responsibilities**

48. 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 Deputy Director, Fabrizio Felloni, and IOE Senior Evaluation Officer, Miguel Torralba.
- Shijie Yang, IOE Evaluation Research Analyst, will provide technical support.

- The Caucasus Research Resource Centre (CRRC), a company selected through competitive processes and tasked with preparing the impact assessment: designing sampling strategy, developing draft questionnaires, training enumerators, undertaking the survey, ensuring quality control in the field, compiling data in electronic form; and selection of advanced statistical analysis techniques and will draft the survey outcome report in collaboration with IOE.
- An international consultant, Mr. Michael Macklin, with solid evaluation skills and good knowledge of Georgian agricultural context, will also be part of the team.

## 7. Timeline

49. During the preparatory phase, a **first reconnaissance mission** was fielded to Georgia in October 2016 in order to make contact with the PMU, and identify national organizations with experience in managing quantitative surveys.

50. 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.

51. 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 Georgia for its comments. The revised and final report will be discussed with the Evaluation Committee.

52. 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 2: Tentative timeline for the evaluation**

<i>Time</i>	<i>Event</i>
September 2016	Selectivity framework for impact evaluation available
October 2016 (2nd week)	Preparatory mission to Georgia
October 2016 (3 <sup>rd</sup> week)	TORs prepared and sent to IFAD Procurement

January 2016	Draft Approach Paper shared with PMD and SKD
January 2016	Methodology developed, including sampling strategy
January 2017 end – April 2017	Field survey designed, conducted (pilot and final), data analysed and draft report on impact evaluation prepared
May 2017 (end)	IOE peer review of draft evaluation report
June 2017 (end)	Draft evaluation report shared with IFAD and government of Georgia
July 2017 (end)	Final report available
August 2017	Learning event in Rome to share key lessons learned
August 2017	Presentation of findings to IFAD and through an in-country event
December 2017	Evaluation report presented to the Evaluation Committee

## **Annex 1 References**

Baker, Judy L. (2000): *Evaluating the Impact of Development Projects on Poverty, A Handbook for Practitioners*, The World Bank, Washington DC.

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.

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.)

NONIE (2009): *Impact Evaluations and Development – NONIE Guidance on Impact Evaluation*, Network Of Networks on Impact Evaluation.

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.

## Annex 2 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<sup>11</sup>.

### 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 2016/2017?	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 CLEs by IOE?	YES
<i>Project status</i>	Did the project implementation end between 1 and 3 years ago?	YES (a little over a year ago)
<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 is undertaking an impact evaluation by the end of 2016?	NO

### Technical criteria

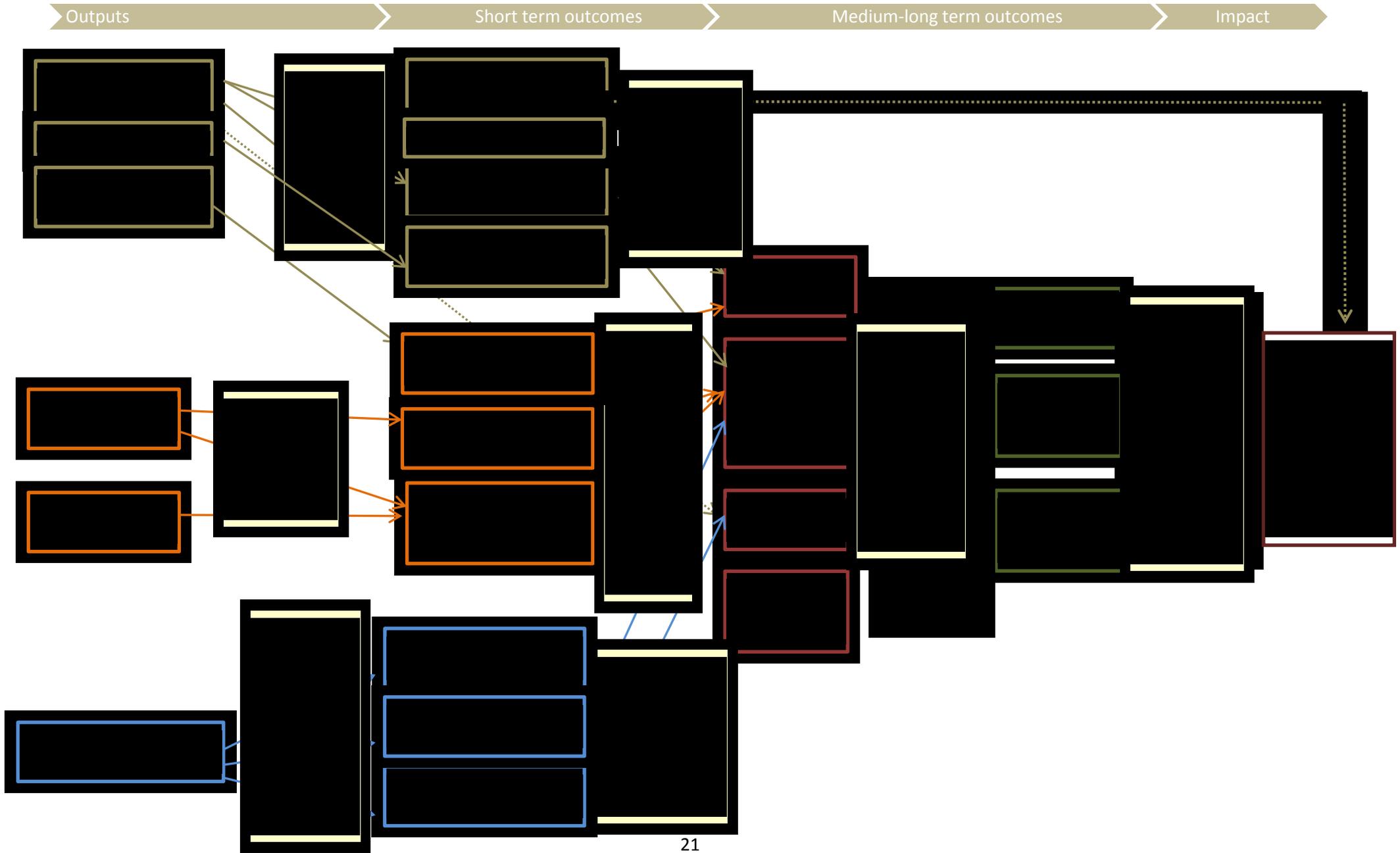
	Is a baseline survey available? What is its quality? Did it include control or comparison groups? Is an electronic database available?	Yes Satisfactory No Yes
	Is a RIMS baseline survey available? <sup>12</sup> What is its quality? Did it include control or comparison groups? Is an electronic database 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?	Yes Satisfactory Yes Yes

<sup>11</sup> 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.

<sup>12</sup> 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>	What is the quality of the PCR including in terms of data and analysis on impact?	To be determined
	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?	Substantive activities started around the 3rd year of the effective date of project
<i>Availability of local technical expertise</i>	Is national technical expertise in quantitative and qualitative data collection and analysis available?	Yes

### Annex 3: Reconstructed Project Intervention Logic



## Annex 4 Logical Framework

Narrative Summary	Impact/Result Indicators	Means of Verification	Assumptions/Risks
<p>Goal</p> <p>The Project goal is to reduce rural poverty in Georgia.</p>	<p>Impact Indicators</p> <ul style="list-style-type: none"> <li>• Reduction in % of rural people living on USD 2/day.</li> <li>• Increase in rural household capital assets.</li> <li>• Reduction in chronic malnutrition among children below 5 years of age.</li> </ul>	<p>DS and LSMS data.</p> <p>Ministry of Health, WHO and World Vision malnutrition surveys.</p> <p>Project M&amp;E database.</p> <p>Mid-term and Completion Assessments.</p>	<p>Political Stability.</p> <p>Macro-economic environment remains conducive to investment, private sector development, and trade.</p> <p>Corruption is contained and its impact on commerce reduced.</p>
<p>Purpose/Objective</p> <p>The Project's objectives are: (i) to increase assets and incomes among actually and potentially economically active poor rural women and men willing to move towards commercially viable agricultural and associated rural enterprises; and (ii) to remove infrastructure bottlenecks which inhibit increasing participation of economically active rural poor in enhanced commercialization of the rural economy.</p>	<p>Result Indicators</p> <ul style="list-style-type: none"> <li>• Value of incremental revenue of primary producers.</li> <li>• Increase in incomes of agro-related employees.</li> <li>• Number of on and off farm new jobs created per USD 1 000 investment through leasing contacts and improved infrastructure.</li> <li>• Increase in public and private commercial investments.</li> <li>• Increase in volume, value, quality and diversity of agro-related trading.</li> </ul>	<p>Mid-term and Completion Assessments.</p> <p>Beneficiaries Assessments.</p> <p>Ad hoc Case Studies.</p> <p>ADPCC and PFI records.</p> <p>Other Government agriculture/trade Data.</p>	<p>Absence of large external economic shocks.</p> <p>No deterioration in internal trade regulations and procedures.</p> <p>Government commitment and understanding of the project.</p> <p>Development and diversification of domestic and international markets.</p>
<p>Outputs from Components</p>	<p>Result Indicators</p>		

<p><b>Support for Rural Leasing</b></p> <p>The recapitalisation and consequent modernisation of Georgian agriculture, specifically among poor smallholders and small and medium agro-related enterprises as a result of the introduction and expansion of rural leasing as a flexible and affordable financial instrument.</p>	<ul style="list-style-type: none"> <li>Type, number and value of leasing contracts.</li> <li>Number and type of PFIs.</li> <li>Production/productivity gains among lessees.</li> <li>Income and capital asset gains among lessees.</li> </ul>	<p>Mid-term and Completion Assessments.</p> <p>Ad hoc Case Studies.</p> <p>ADPCC and PFI records.</p> <p>Export/import Statistics.</p>	<p>No major adverse developments in financial sector stability in Georgia.</p> <p>External markets for Georgian agro-products diversified or reopened after the 2008 conflict and import substitution policies in place.</p>
<p><b>Narrative Summary</b></p>	<p>Impact/Result Indicators</p>	<p>Means of Verification</p>	<p>Assumptions/Risks</p>
<p><b>Outputs from Components</b></p>	<p>2nd level Result Indicators</p>		
<p><b>Small-scale Rural Infrastructure (SSRI)</b></p>	<ul style="list-style-type: none"> <li>Area of rehabilitated or established irrigated land.</li> <li>Water delivered compared to water requested.</li> <li>Km of rural roads rehabilitated.</li> <li>Number and type of other ASP-supported infrastructure.</li> <li>Number of functioning infrastructures after 3 years.</li> <li>Number and type of created or expanded businesses as a result of developed infrastructure.</li> <li>Incremental annual value of revenue of farmers/enterprises served by infrastructure at establishment and after 3 years.</li> <li>Value of villagers' contribution in support to infrastructure projects.</li> <li>Number of beneficiaries by type of rehabilitated/developed infrastructure.</li> </ul>	<p>Contractors' reports.</p> <p>Mid-term and completion assessment.</p> <p>Case studies.</p> <p>PIU/ADPCC field visits and records.</p> <p>Business plans and subsequent records.</p> <p>Beneficiary focus group discussions.</p>	<p>Transparent criteria applied for awarding of contracts.</p> <p>Interest of Government and potential contracted processors to participate in the project.</p>

## Annex 5 Evaluation Framework

<b>Evaluation criteria</b>	<b>Key evaluation questions</b>	<b>Performance Indicators</b>	<b>Data Sources</b>
Rural poverty impact	Was there an improvement in the socio-economic situation of beneficiaries? Were beneficiaries below poverty line lifted out of poverty? Was resilience of beneficiaries to economic shocks improved? 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?	Human Development Index (as far as possible) 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? Did the saving capacity of beneficiaries improve as compared to project baseline? How are those savings utilized? 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 Increased savings and credit	Household survey Focus Group Discussion
<i>b) human and social capital and empowerment</i>	To what extent did the project contribute 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 Access to education and health facilities Access to safe source of drinking water	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 the project affect the quantity and reliability of irrigation water provided to farmers? 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 cost/value of production, average area under production, average yields per hectare, average milk production per household.	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? To what extent and how did the project contribute to improve the resilience of beneficiaries to environmental shocks?	Soil and water management, sustained production under climate variability.	Household survey Focus Group Discussion
<b>Project performance</b>			
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?	Proxy indicators of relevance	Desk review Key Informant Interviews

	<p>Was it based on development approaches tailored to the context?</p> <p>Did the project have an exit strategy at design?</p>		
Effectiveness	<p>Was the project targeting approach effective?</p> <p>What was the project outreach at completion?</p> <p>Did the project meet its objectives?</p> <p>For instance, did the project affect the quantity and reliability of irrigation water provided to farmers? Did farmers adopt new agricultural practices as a result of the project? Did rehabilitating bridges improve access to markets and social infrastructure? Did leasing create increased secondary effects in the form of poor farmers benefiting from increased demand for their commodities from leasing enterprises and additional employment for poor rural people? Did the above result in increased crop production and/or crop diversification?</p>		<p>Desk review</p> <p>Household survey</p> <p>Focus Group Discussion</p> <p>Secondary data</p> <p>Key Informant Interviews</p>
Efficiency	<p>How economically resources and inputs were converted into results?</p> <p>Were the project effects large enough to justify its costs? (Economic Rate of Return)?</p> <p>What was the cost of the project as compared to projects supported by other donors in the country?</p> <p>What was the time lag between approval and loan effectiveness? What was the budget utilization at completion?</p> <p>Were the funds from IFAD and other partners made available in a timely manner?</p> <p>What are the project management costs at completion? And compared to other similar projects?</p>		<p>Desk review</p> <p>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?</p> <p>Is there evidence that the infrastructure investments will be sustained after rehabilitation was completed?</p>		<p>Household survey</p> <p>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?</p> <p>Did the project strengthen women's agencies – their decision-making role in community affairs and representation in local institutions?</p> <p>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</p> <p>Focus Group Discussion</p>
Innovation and scaling-up	<p>To what extent did the project introduce innovative approaches that have been scaled-up by the government and others?</p> <p>To what extent did the project learn from past experience and inform the design of new projects?</p>		<p>Desk review</p> <p>Key Informant Interviews</p>