June 2010

83157

Strengthening National Comprehensive Agricultural Public Expenditure in Sub-Saharan Africa

Expenditure Component Impact Evaluation

Template Terms of Reference

web page: www.worldbank.org/afr/agperprogram

email contact address: agperprogram@worldbank.org

1. Background

The programme "*Strengthening National Comprehensive Agricultural Public Expenditure in Sub-Saharan Africa*" financed by the Bill and Melinda Gates Foundation and implemented by The World Bank, seeks to improve the impact of scarce public resources spent by Sub-Saharan African governments on agricultural sector development activities, thereby improving the welfare of predominantly poor rural populations. It operates in the context of the Comprehensive Africa Agriculture Development Programme (CAADP) of the African Union's (AU) New Partnership for Africa's Development (NEPAD) which encourages governments and development partners (DPs) to target public expenditure on the agriculture sector as the most effective way of stimulating growth in the sector, thereby reducing hunger and poverty.

The overall objective of the programme is, through providing analytical support, to promote the articulation and implementation of strengthened national comprehensive agricultural public expenditure programmes so as to build consensus for increased levels of public expenditure in the sector in Sub-Saharan Africa, and to enhance its efficiency, effectiveness and equity.

The programme is intended to provide evidence-based recommendations that will address, *inter alia*, budgetary planning, budget execution, accountability, and stimulate larger donor resource allocations, enhanced harmonization and alignment of resources behind national strategies, the creation of a reliable data base, and more effective intra- and inter-sectoral coordination. In the specific context of CAADP, the study will focus on: the level of expenditure on agriculture, with particular reference to the explicit target by African Heads of State in the Maputo Declaration to allocate 10% of national budgets to the sector; the composition and priorities of expenditure with respect to stated national strategies, evidence of impact, sustainability and absorptive capacity; and, budget planning and implementation so as to strengthen public financial management in general and in particular budget coherence, outputs, outcomes and supporting mechanisms such as procurement and audit.

Template Terms of Reference

Two different levels of analytical support will be provided through the programme: (a) to conduct a basic agriculture sector public expenditure review (PER) in countries where this work has not already been undertaken recently, and (b) to carry out specialised public expenditure analyses in situations where an agriculture sector PER already exists. The three specialised studies are: the development of a sectoral medium term expenditure framework (MTEF); public expenditure tracking surveys (PETS); and, expenditure component impact evaluation (the subject of these terms of reference). The existence of an agriculture sector PER is the prerequisite for conducting the other specialised studies. This should provide the essential understanding of the structure of public support for the sector and of the flows of funds that result. The public expenditure tracking survey (PETS) which examines outputs of public programmes, and the expenditure component impact evaluation study which looks at outcomes, are both backward looking exercises designed to inform policy and programme design. They both feed into sector investment programmes as part of the sector MTEF which is forward looking. The output of the analytical support will be country-specific reports that have been prepared jointly by government staff from the Ministries of Agriculture and of Finance together with external technical assistance.

The purpose of preparing template terms of reference (TOR) is to provide a clear framework which defines the scope, methodology and processes that should be adopted when carrying out each country study,

whilst allowing flexibility for the task to be tailored to the specific needs, data availability and analytical capacities that exist in each country. The terms of reference serve as a checklist of items that should be covered in each study together with an indication of the level of detail that should be attempted and possible data sources and approach. The template framework aims to ensure that, as far as possible, data and analyses can be compared across different countries.

The template TORs draw heavily for their methodology upon the "*Practitioners' Toolkit for Agriculture Public Expenditure Analysis*" (APEA) put together by World Bank and DFID[1] and the Bank's "*Public Expenditure Management Handbook*"[2].

2. Scope

The purpose of the expenditure component impact evaluation specialised study is to assist selected governments in sub-Saharan Africa (SSA) to enhance the outcomes and sustainability of their agriculture sector programming and budgeting and the efficiency and effectiveness of their public expenditure in the sector. Through carrying out evaluations to assess the outcomes or impact of specific areas of public expenditure in the agriculture sector, this study should strengthen the evidence base for policy making and scaling-up programmes as well as creating a "management-for-results" environment and enhancing the efficiency of budget planning. Its results can also be used to identify corrective actions in existing programmes to ensure that both current and future investment in key subsectors generate the expected and sustainable outcomes.

Impact evaluation is one of several different ways in which a government (or donor) may compare *ex ante* investment feasibility projections regarding the impact of a specific public investment with actual outcomes or *ex post* impact. An impact evaluation complements tools such as implementation completion reports (ICRs) prepared by governments and lenders, and public expenditure tracking surveys (PETS) of the type that will be undertaken as part of this programme, that are more concerned with service delivery outputs. A formal impact evaluation demands greater analytical rigour than some of the more descriptive (Basic AgPER or an ICR, for example) or diagnostic exercises (such as a PETS) that constitute part of this overall programme.

Scale

The scale and complexity of impact evaluations vary greatly. At the simplest level, they may be used to determine the impact of one specific aspect of a large public investment or project. At the other extreme, they may attempt to assess the impact on the economy as a whole – the impact on growth and poverty reduction, for example – of an entire sector public investment programme. Under this programme, the emphasis will be on conducting rather narrowly focused impact evaluations such as for a component or sub-component that a basic sector PER has indicated is an important part of the overall investment programme and for which there is scope for improved outcomes, so that the limited resources available can produce credible results.

[1] Practitioners' Toolkit for Agriculture Public Expenditure Analysis, World Bank and DFID (draft)
[2] "*Public Expenditure Management Handbook*", World Bank, 1998

Survey Focus

As outlined above the impact evaluations conducted under this programme will be limited in scope. Within the agriculture sector, each evaluation is likely to cover just one of the following:

- Extension and training
- Agricultural research
- Input supply interventions
- Support for marketing and value chains
- Infrastructure development (such as: irrigation, feeder roads, market information systems, post-harvest handling)
- Support for a specific commodity or product group (such as specific crops, livestock, fisheries, or forestry products)
- Institutional issues such as access to land or credit

Country Selection

In order to make a significant contribution to the effectiveness of public expenditure, an impact evaluation should be carried out in countries which meet the following criteria:

- Have a sufficiently sound level of public expenditure management (PEM) to ensure that the use of funds can be ascertained in some detail
- Have a sufficiently articulated sectoral evaluation system and programme monitoring and evaluation (M&E) framework in place to provide a data base which allows outputs to be quantified and at least part of the impact of programmes on households to be measured
- Have recently completed an agriculture sector PER (possibly the Basic Ag PER which forms part of this overall programme), so that there is a clear understanding of the flow of funds within the sector institutions and between different administrative levels (flows to decentralised units)
- Where governments are prepared to allow the data and analysis carried out to be made available online for the purpose of training the staff of MOAs and MoFs in other countries of the region

Institutional Scope

The line ministry (usually the Ministry of Agriculture) or related agencies that are responsible for the programme for which the impact is to be evaluated can be selected as the partner for an impact evaluation. The choice of thematic focus would determine the responsible institution.

Implementation

The exercise is intended to provide evidence-based insights into the impact of public investments to high level policy and decision makers. The results of such a survey would be used, for example, to justify a proposal to scale-up an existing programme or to cut back or modify a programme which is shown not to be having the intended impact. When major strategic choices are to be made on the basis of such an evaluation, it is important to ensure that the sector line ministry or agency concerned is fully involved in

the exercise. However, care should be taken to avoid undue influence on the part of those engaged in the programme being evaluated on the way the study is undertaken. For this reason, a senior level, multi-stakeholder steering group (SG) would provide oversight of the exercise and facilitate access to high level decision-makers.

The programme will be conducted by a TA team comprising one international expert with the necessary analytical skills, and two national experts with knowledge of the sector and evaluation methodology. The national experts will bring country and specific sub-sector knowledge as well as analytical skills and will share responsibility with the international expert for survey design, data gathering, analysis and report writing. The specific outputs of the TA team are listed in section 6 below.

The timing and length of the survey would depend upon its focus and the extent to which there are important seasonal characteristics that would need to be captured. As explained below, the duration of the study will depend upon the complexity of the task. As a minimum, it is likely that a period of five months would be required to complete a comparatively simple evaluation. For more complex situations, nine months or longer might be needed.

Donor Engagement

The active engagement of the donor working group (DWG) in the agriculture sector in the identification of or endorsing the thematic area for the evaluation, its design and ultimately the discussion of results is important, not least because donors will almost always have played an important part in designing and financing the programme being evaluated. Donors increasingly demand evaluations of the impact of the programmes they have financed, and it is especially important that this be carried out in a rigorous way when new phases or scaling-up are being considered. In addition, the involvement of the DWG will:

- Facilitate launching discussions with government on possible priority areas for an evaluation before the TA team in engaged
- Ensure that DPs have the opportunity to express their concerns over certain public investment programmes
- Ensure that once the evaluation is completed, the DPs are able to endorse the results and facilitate the dialogue with government on recommended strategies for scaling-up or changing investment programmes
- In cases where the DWG is strongly committed to an evaluation which is likely to be “complex” in nature, to facilitate accessing supplementary financing of the task from DWG members

3. Methodology

The scale and complexity of the impact evaluation that will be undertaken is likely to vary greatly from country to country. This will reflect the intrinsic complexity of the expenditure component selected as a focus as well as the extent to which existing data are considered to be adequate and reliable. These factors, *inter alia*, will determine the scale of data gathering required, the length of time that will be needed for the survey and, of course, the cost of the exercise. For this reason it is difficult to propose a standard methodology. The table below summarises the likely characteristics of a comparatively “simple” impact evaluation with those of a more “complex” survey.

	Simple Evaluation	Complex Evaluation
Survey focus	• Small country • Comparatively new service or programme • Largely undifferentiated population with respect to topic • Centrally managed	• Large country • Large-scale, long-term service • Need for careful stratification of sample • Many administrative layers • Decentralised responsibilities
Data availability	• Good M&E system in place • Baseline undertaken • Target and non-target population data available	• Poor or non-existent M&E system • No baseline data available • Data do not address survey issues
Data gathering	• TA team • Small counterpart team	• TA team • Counterparts • Interviewers/enumerators
Additional costs	• Small – can be accommodated within existing budget	• Substantial costs for hiring and mobilising interviewers and enumerators

The main factors determining the methodology to be employed for the impact evaluation include:

- **Analytical skills**: Unlike other parts of this programme, an impact evaluation demands analytical skills, such as modelling and regression analysis that are often in short supply in Ministries of Agriculture. This will probably mean that most of the analysis will be undertaken directly by the TA team and this will determine the type of TA support provided.
- **Available data**: There will be heavy reliance upon existing sources of data and the availability and quality of these data will determine the scope of the evaluation.
- **M&E**: Much of the data will come from M&E systems in place and if these do not generate evidence-based results, it will be difficult to carry out a thorough evaluation without substantial additional data gathering.
- **Limited resources**: The limited resources available for the survey means that:
    - The topic of the evaluation should be carefully targeted
    - Farm level data collection should be kept to a minimum
    - Existing sources of data should be used for the evaluation, including beneficiary impact assessments, household income and expenditure surveys, standard of living surveys, poverty assessments, crop and livestock surveys etc
- **Sub-sector selection**: The methodology will have to be tailored to the specific sub-sector or type of expenditure component selected

Setting up an Impact Evaluation

There are four steps that should be taken when setting up an impact evaluation:

1. Selection of the evaluation topic
2. Assessment of data and information sources
3. Supplementary data gathering
4. Data analysis, reporting and dissemination

1. Selection of Evaluation Topic

The selection of the topic or theme for the impact evaluation must be carried out in close collaboration with the line ministry concerned and with donors. Broad agreement on the focus area for the impact evaluation should have been reached prior to the engagement of the TA team since the sub-sector selected might well demand specific expertise.

Selection criteria: As outlined above, there are a limited number of potential key areas for an evaluation but the opportunity should also be taken to add cross-cutting dimensions. This might involve looking at the gender dimensions of irrigation development, for example, or access to credit in a specific value chain. In countries where CAADP has reached the compact stage, the background material contained in the "*Stocktaking and Institutional Analysis*" which forms part of the CAADP Country Roundtable process, could provide useful insights into critical areas which would merit further evaluation.

The criteria to apply for the selection of a topic or theme could include:

- The programme Is a core part of the sector
- Its share of total public expenditure in the sector is large (say, >20%)
- There is scope for replication or scaling-up
- It has innovative features the impact of which need substantiation
- It is a substantial programme but there are questions over its impact

The selection of a topic for the evaluation must be accompanied by a thorough understanding of the ways in which specific public investments are expected to achieve the expected impacts. At this stage, the TA team and its counterparts should propose clearly-defined research hypotheses and also set out one or more "results chains" linking public investment to impacts.

Indicators: Part of the preparatory process will also involve deciding on the key indicators of impact that will be measured. The starting point for this will be to examine the development objectives of the programme and, if one exists, the logframe which was part of the programme design and appraisal. Although a logframe will have been prepared for programmes financed by development assistance, this might not be the case for a domestically-funded programme or one in which only certain components were externally financed. The types of evaluation that are likely to be undertaken as part of this programme will be interested in assessing five types of impact:

- **Return to investment**: calculating *ex post* economic rates of return or net present value using cost-benefit analysis; impact on leveraging private sector investment
- **Household level impact**: calculating incremental incomes; financial rate of return; analysis of profitability of farm enterprises; enhanced access to markets; gender dimensions
- **Technical impact**: impact on crop or livestock productivity
- **Institutional impact**: enhanced skills and capacity in the sector institutions; improved governance of the sector;
- **Sustainability**: assessment of recurrent costs of the programme; revenue generation; beneficiary commitment to processes or technologies

A difficult part of the analysis will be to devise a way to adequately filter out or control for the "noise" in the data when attempting to determine cause and effect relationships in impacts. For example, seasonal factors affecting input and output prices, market variability, good and bad rainy seasons, and so on. In

addition, factors beyond the control of the programme such as natural disasters, and shortcomings in implementation such as lack of counterpart funds, poor governance and budget related issues, need to be identified.

2. Assessment of Data and Information Sources

An assessment must be made of both primary and secondary data and information sources, looking especially at data in the sub-sectors and areas of interest for an evaluation. Primary data are most likely to come from M&E systems associated with specific projects or programmes. The main features to assess, depending on the type of topic are:

- **Time series**: preferably time series data for the length of the programme, but a minimum of two years
- **Level of detail**: data down to the lowest administrative unit and/or to farmer group level, covering all areas in which the programme operates
- **Quality**: internal consistency and absence of blank data sets
- **Baseline**: M&E data sets which do not have a fully documented and plausible baseline should be treated cautiously
- **Raw data**: availability of raw data in electronic form for data verification and, if not fully tabulated and analysed, for tailored analysis

Secondary information and data sources are most likely to be useful for cross-checking conclusions drawn from primary data. For example, evaluations or beneficiary impact assessments undertaken on similar or related topics within the sector can provide cross-checks on consistency.

3. Supplementary Data Gathering

An impact evaluation is obliged to depend heavily upon existing sources of data. However, if there are serious gaps in the data needed to test a hypothesis this will need to be filled by generating new data in the field. In practice, even the best M&E system is unlikely to have gathered data that will directly address the types of questions posed by an impact evaluation. However, such supplementary data gathering should be minimised, highly focused and kept as small as is compatible with achieving credible results. Depending upon the type of topic being evaluated, comprehensive data at the household level are most likely to be lacking and some form of beneficiary impact assessment (BIA) might be required. The main steps in gathering additional data would be:

- **Sampling and questionnaire design:** Both quantitative and qualitative data can be collected. Although desirable, it is unlikely that the survey instruments employed can be standardised across countries in view of the likely diversity in focus topics and the different capacities to support the exercise at country level. Additional data collected by a targeted survey should be carried out on a random sample basis and be stratified and framed carefully so that the survey results are credible from a statistical point of view. The sample should reflect the typical variability that exists (for example, well served areas, badly served areas and average areas) as well as factors such as remoteness, rich and poor areas, and so on. The effectiveness of drawing a random sample will depend upon the existence of clear data on population distribution and household characteristics.
- **Prepare draft questionnaire**: The questionnaire should be kept as simple as possible and tailored to the different administrative levels involved and the target interviewees including, for example,

central government staff, local government staff, farmer groups and farmers. The questionnaire should be field tested and administered by experienced enumerators or interviewers.

- **Implementation**: For a "simple" impact evaluation it is likely that most of the survey can be conducted by the TA team and a small team of counterparts. For more "complex" evaluations will require the recruitment of dedicated interviewers or enumerators. This will involve identifying enumerators/interviewers and, if necessary, giving them the necessary training; field testing of the draft questionnaire on a pilot scale to allow adjustments and refinements to be made before launching the survey; data collection, with adequate monitoring and supervision of the task; data entry and cleaning; if a known and trusted locally-used system is not available, the Census and Survey Processing System (CSPro)[3] from the US Census Bureau is available as a free download.
- **Case studies**: Valuable insights into specific aspects of an evaluation can be validated and illustrated by the use of case studies, especially if high quality data are lacking.

4. Data Analysis, Reporting and Dissemination

Data analysis: The process of data analysis will be carried out by the TA team in close collaboration with staff from the partner line ministry. On-the-job training would be part of the TA team's responsibility, but this will be limited mainly to ways of managing the impact evaluation process so that this can become a routine part of the MoA's programming process.

Reporting: The first report to be produced would normally be a summary of the data analysis which contains the initial conclusions of the evaluation. This summary report should be discussed with the line sector ministry or agency concerned before producing a full report. This will normally include the full analysis of the data together with conclusions on the impact of the programme being evaluated.

Dissemination: The dissemination of the impact evaluation report would have to be agreed with the ministry or agency concerned. The report should be distributed to senior government officials, politicians and the DPs and should also be discussed at a workshop for all the key stakeholders.

Recommendations

In addition to the conclusions emerging from the impact evaluation, the TA team should also make recommendations that might go beyond the scope of the report itself including:

- Where data availability or quality have proved to be an issue, the need to introduce a functional or enhanced M&E framework for the sector
- The creation of institutional and sectoral data systems to enable better analysis to be conducted in the future

[3] Version 4.0 free download, only in English at: http://www.census.gov/ipc/www/cspro

4. Sources of Data and Information

An impact evaluation utilises existing data as well as generating specific new data in the agriculture sector. The key to a successful and useful impact evaluation is to start with a sound understanding of the sector, public expenditures and the main programmes that comprise the public investment. The main sources will be:

Official sources:

- Programme and project design/appraisal documents
- Electronic data from MoA and MoF
- Published reports and statistics from MoA and associated agencies such as the Ministry of Trade and/or Commerce

Secondary sources:

- MoA studies and reports
- Project mid-term reviews, ICRs
- Donor reports on programme implementation
- CAADP Country Roundtable reports
- Poverty assessment reports
- Household income and expenditure surveys
- Crop and livestock survey reports and data
- Focus group discussions with project managers and teams
- Donor evaluation reports
- Beneficiary impact assessments

5. Process

A fully participatory approach is to be adopted for the task. The most important first step in starting the exercise is to establish a good working partnership with the key stakeholders: MoA, MoF, the donor working group (DWG) for agriculture, CAADP Focal Point, parliamentarians, representatives of the private sector, and CSOs. In the case of MoA, technical level counterparts must be assigned to work with the technical assistance team (international and national) on the task on an intermittent basis for a period of five months.

The study process will comprise three phases, each of which will require the active engagement of the TA team:

1. Preparatory Phase

Stakeholder briefing: All key stakeholders[4] must be briefed on the purpose and proposed outputs of the exercise and agree on the main milestones and the timeline. Agreement should also be reached upon the thematic focus of the evaluation.

Inception workshop: A formal workshop for all key stakeholders should be scheduled within two weeks of the commencement of the study to present and discuss an Inception Report that includes:

- The revised terms of reference for the study
- Proposed areas as the topic for the evaluation
- Data and information sources
- Formulation of research hypothesis
- Queries and outstanding issues with respect to the evaluation
- Agreed milestones and timeline for implementation
- Designation of counterpart staff

The workshop should solicit the support of government officials in facilitating access to data (including the agriculture sector PER) and reports on different aspects of sector programmes, and from DPs (through the DWG) on programme evaluations that have been conducted.

Setting up a steering group: A senior level steering group (SG) should be set up to oversee the evaluation. It should comprise sector specialists in agriculture, staff responsible for the sector within the MoF, representatives of CSOs and the private sector and a representative from the DWG. The SG would:

- Review the Inception Report
- Agree on the impact evaluation topic
- Review the summary report
- Review the final report
- Facilitate dialogue with senior levels of government on the conclusions of the evaluation

2. Implementation Phase

Survey implementation: Including the design of the sampling frame and questionnaire for supplementary data gathering, field testing and actual implementation of the evaluation survey.

Technical workshop: A technical workshop would be held after approximately two months to review progress with data assembly, outlining the types of analysis to be undertaken and arrangements for any supplementary data gathering.

Data analysis: Data entry and analysis would commence as soon as the first survey data are produced.

3. Reporting Phase

Preparation of draft summary report: The summary report would be prepared as soon as data analysis is completed.

[4] The CAADP Country Roundtable consultative process provides a possible scope for relevant stakeholder consultations.

Draft report workshop: A formal workshop should be scheduled after four months to discuss the key conclusions of the evaluation, presented in a summary report. The TA team would be primarily responsible for the presentation of the draft report

Final Report: The final report would be prepared reflecting the discussions at the draft report workshop and comments from government and the DWG.

6. TA Team Outputs, Reports and Database

TA Team Outputs

The TA team, in close collaboration with counterparts from MoA will be responsible for the following tasks:

- Discussing with government and donors the topic for the evaluation
- Helping to refine the research hypothesis
- Preparing a "results chain"
- Designing the supplementary data gathering
- Carrying out data analysis
- Carrying out on-the-job training in impact evaluation methodology and management
- Preparing a summary report
- Preparing a full evaluation report
- Designing and managing the inception, technical and draft report workshops
- Liaising with the DWG
- Formulating recommendations for enhancing sector M&E systems and the utilisation of evaluation results
- Ensuring that all data utilised during the evaluation is assembled in a form that can be put online

In addition, the TA team would play an active part in refining, through discussions with government and DWG, the focus area for the impact evaluation survey.

Reports

Following consultation with all the main stakeholders, the TA team will prepare the following reports:

- **Inception report**: within two weeks, which presents the revised terms of reference for the evaluation, raises issues and defines the topic for the evaluation
- **Summary report**: upon completion of the initial data analysis, which presents the main results of the evaluation
- **Full evaluation report**: within one month of the completion of the initial analysis, which includes all data analysis and presents the conclusions of the evaluation
- **Final report**: within five months, incorporating the TA team's broad recommendations on the M&E system and framework for using the evaluation results

Database

During the assignment, the TA team in collaboration with MoA and MoF will establish a database including background documents and sources, evaluation methodology, data sampling frame and questionnaires, survey data, analytical working papers and analyses that can be put online. This database will be used for capacity building purposes as part of a joint learning activity under the overall public expenditure programme.

7. Timeline

The overall timeline for conducting a component impact evaluation in the agriculture sector is at least five months (see Fig. 1) for a "simple" evaluation, but could be nine months or more for more "complex" situations. In terms of the timeline chart, additional time is most likely to be required for Phase 2, reflecting the scale of data gathering required.

8. Resources

The programme will finance a consultant technical assistance (TA) team comprising one senior international expert with methodological expertise in impact evaluation and cross-country experience, plus two national experts. The experts will each devote 12 weeks of work to the exercise within the overall time frame required. The international expert will make three visits to the country, corresponding to the three phases of implementation. The international expert will be consulted in the selection of the national expert.

The Ministry of Agriculture or other concerned sector agency will be requested to provide at least one full-time counterpart or team equivalent for each consultant. Think tanks, universities and other local contributors to the exercise will not be remunerated.

In the case of more "complex" situations where substantial supplementary data gathering must be undertaken, the costs of recruiting and mobilising interviewers and enumerators will need to be financed outside the programme's budget, possibly through the DWG.

Annexe - Timeframe