Global Donor Platform
for Rural Development

Tracking results in agriculture and rural development in less-than-ideal conditions

A sourcebook of indicators for monitoring and evaluation

Tracking results in agriculture and rural development in less-than-ideal conditions

A sourcebook of indicators for monitoring and evaluation

Published by

GLOBAL DONOR PLATFORM FOR RURAL DEVELOPMENT

FOOD AND AGRICULTURE ORGANIZATION
OF THE UNITED NATIONS

THE WORLD BANK

This publication can be downloaded from the following websites:
1. www.donorplatform.org
2. www.worldbank.org
3. www.fao.org

The request for hard copies should be sent to any of the co-publishers at addresses below:

1. Secretariat of the Global Donor Platform for Rural Development,
Dahlmannstrasse 4, 53113 Bonn, Germany
Fax: +49 (0) 228 24 934 155
e-mail: secretariat@donorplatform.org

2. The World Bank
Sector Manager, Agriculture and Rural Development Department,
The World Bank,
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A
Fax: +1 202 522 3308

3. FAO Statistics Division
Viale delle Terme di Caracalla, 00153 Roma, Italy
Fax: +39 06 570 55615
e-mail: ESS-Registry@fao.org

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations, of the World Bank or of the GDPRD concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO, the World Bank or the GDPRD in preference to others of a similar nature that are not mentioned. The views expressed herein are those of the authors and do not necessarily represent those of FAO, the World Bank or the GDPRD

ISBN 978-92-5-106082-7

All rights reserved. Reproduction and dissemination of material in this information product for educational or other non-commercial purposes are authorized without any prior written permission from the copyright holders provided the source is fully acknowledged. Reproduction of material in this information product for resale or other commercial purposes is prohibited without written permission of the copyright holders. Applications for such permission should be addressed to the Chief, Electronic Publishing Policy and Support Branch, Communication Division, FAO, Viale delle Terme di Caracalla, 00100 Rome, Italy or by e-mail to copyright@fao.org.

© GDPRD, FAO and World Bank, 2008

Contents

ACKNOWLEDGEMENTS vii

ACRONYMS AND ABBREVIATIONS ix

PREFACE xi

EXECUTIVE SUMMARY xiii

CHAPTER 1 - THE EVOLUTION OF M&E IN DEVELOPMENT 1
- What is M&E? 1
- Who are the users? 5
- How M&E has evolved? 7

CHAPTER 2 - THE ANALYTICAL FRAMEWORK 13
- Thinking logically about indicators 13
- Monitoring performance (inputs and outputs) 17
- Measuring results (outcomes and impact) 20
- A core set of priority indicators for ARD programmes 33

CHAPTER 3 - THE DATA FRAMEWORK 39
- The tools 40
- Applying the tools for M&E analysis 54
- Strengthening National Statistical System capacity 58

CHAPTER 4 - THE INSTITUTIONAL FRAMEWORK 63
- The M&E framework 63
- The statistics framework 70
- The international framework 73
- The role of development partners 75

CHAPTER 5 - SETTING UP AN M&E STRATEGY IN AGRICULTURE AND RURAL DEVELOPMENT 79
- Step 1: Assessment and diagnosis 80
- Step 2: Review of indicators 80
- Step 3: Review of current data, sources and gaps 81
- Step 4: Development of action plans 81
- Step 5: Review of resource requirements 81
- Step 6: Monitoring the performance of the M&E action plan 83

BIBLIOGRAPHY **85**

ANNEX 1: A LIST OF CORE INDICATORS **89**

ANNEX 2: COUNTRY CASE STUDIES - SUMMARY OF COUNTRY STUDIES AND OF ARD INDICATORS CURRENTLY IN USE IN EACH COUNTRY **111**

Part 1 – Country studies 111

Part 2 – ARD indicators in use in each country 124

ANNEX 3: M&E CAPACITY ASSESSMENT SCORECARD **132**

LIST OF BOXES

Box 1.	Definitions of Monitoring and Evaluation: 1984-2002	2
Box 2.	How M&E findings help governments and stakeholders?	6
Box 3.	Poverty Reduction Strategy Papers (PRSPs)	11
Box 4.	The national management information system for local government reform of the United Republic of Tanzania	19
Box 5.	Characteristics of different classes of indicators	21
Box 6.	Adaptation of research and extension service delivery indicators (access, use and satisfaction) to the new Technology Transfer Paradigm	24
Box 7.	Detecting a trend in maize yields	28
Box 8.	List of priority indicators	35
Box 9.	Cambodia's two-tiered system	36
Box 10.	Tools for measuring results: surveys vs. non-formal appraisal methods	43
Box 11.	Core Welfare Indicators Questionnaire (CWIQ): a survey instrument for collecting service delivery indicators	48
Box 12.	Nigeria's community service delivery survey	52
Box 13.	Comparison of key features of different surveys	54
Box 14.	How do we know if a Poverty Reduction Strategy is effective?	64
Box 15.	The M&E system of Cambodia's Ministry of Agriculture, Forestry and Fisheries	66
Box 16.	The Poverty Reduction Strategy Monitoring Master Plan (MUKUKUTA) of the United Republic of Tanzania	67
Box 17.	M&E Technical Committee – sample Terms of Reference	68
Box 18.	National Planning, Monitoring and Evaluation (PM&E) Workshop in Nigeria	69
Box 19.	Nicaragua – Linking the M&E activities more closely with the National Statistical System	71
Box 20.	Senegal's Reformed National Statistical System	72
Box 21.	Agriculture and the Millennium Development Goals	74
Box 22.	National Statistical Development Strategy essentials	76
Box 23.	A results chain for building an M&E system	82

Acknowledgements

This Sourcebook was prepared by a joint team of staff from the World Bank and the Food and Agriculture Organization of the United Nations (FAO), led by Nwanze Okidegbe (World Bank) and consisting of Tim Marchant (principal consultant); Hiek Som, Naman Keita, Mukesh K. Srivastava and Gladys Moreno-Garcia (FAO Statistics Division); and Sanjiva Cooke, Graham Eele, Richard Harris and Diana Masone (World Bank). Research assistance was provided by Patrice Wadja, Zena Angesom and Maria Rey de Arce (consultants). The editing was done by Barbara Hall and the layout and design, by Marianne Sinko.

The team is grateful for the comments received from the following peer reviewers: Fred Vogel, Gershon Feder, Stephen Mink (World Bank); Neela Gangadharan (FAO); Focal Points of the Global Donor Platform for Rural Development (GRPRD) and the Wye Group; Krijn Poppe (Agricultural Economics Research Institute); and Jan Karlsson (United Nations Economic Commission for Europe, retired). Comments on the initial menu of indicators were gratefully received at an expert group meeting held in Washington D.C., United States of America, especially those of Susan Offut (United States Congressional Service), Haluk Kasnakoglu (FAO, retired), and Misha Belkindas and Haeduck Lee (World Bank).

Two consultants, Miguel Galmes and Generoso De Guzman, participated in the preparation and implementation of in-country validations. In each country, the validation was led by a team of two national consultants under the technical supervision and support of FAO staff. The following national consultants conducted the country validation studies: Chantum Chea and Monthivuth Ker (Cambodia); Patrick Dumazert and Karla Arriola (Nicaragua); Yahaya Husseini and Frederick Okoukoni (Nigeria); Aboubacry Demba Lom and Mamadou Wane (Senegal) and Evelyne A. Lazaro and Radegunda Maro (United Republic of Tanzania). A number of national experts from various government departments in the countries dealing with agriculture and rural development (ARD) made valuable contributions to the work. In all these countries, National Statistical Offices (NSOs) and Planning Departments played a key role. A number of development partners representing the donor community in the countries took active part in National Validation Seminars. The team is also grateful for the valuable administrative assistance provided by FAO representatives in the countries and their staff.

The team thanks Juergen Voegele, Mark Cackler, Christopher Delgado and the Management Team of the Agriculture and Rural Development Department of the World Bank, the relevant Technical Divisions of FAO and the Secretariat of the Global Donor Platform for Rural Development for their support and inputs.

Acronyms and abbreviations

ARD	Agriculture and Rural Development
ASDP	Agricultural Sector Development Programme (the United Republic of Tanzania)
CMDG	Cambodian Millennium Development Goal
CWIQ	Core Welfare Indicators Questionnaire
DAC	Development Assistance Committee
DFID	UK Department for International Development
FAO	Food and Agriculture Organization of the United Nations
GDDS	General Data Dissemination System
GDP	Gross domestic product
GDPRD	Global Donor Platform for Rural Development
GPS	Global Positioning System
HBS	Household Budget Survey
HIPC	Highly Indebted Poor Countries
I&D	Irrigation and drainage
ICT	Information and Communications Technology
IDA	International Development Agency
IFAD	International Fund for Agricultural Development
LGA	Local government authority
LGRP	Local Government Reform Programme (the United Republic of Tanzania)
LSMS	Living Standards Measurement Study
M&E	Monitoring and Evaluation
MDG	Millennium Development Goal
MIS	Management Information System
MUKUKUTA	Poverty Reduction Strategy Monitoring Master Plan of the United Republic of Tanzania
MoP	Ministry of Planning (Cambodia)
NGO	Non-governmental organization
NIS	National Institute of Statistics (Cambodia)
NSDP	National Strategic Development Plan (Cambodia)
NSDS	National Strategies for the Development of Statistics
NSO	National Statistical Office
NSS	National Statistical System
OECD	Organisation for Economic Co-operation and Development
PARIS 21	Partnership in Statistics for Development in the 21st century

PETS	Public Expenditure Tracking System
PPP	Purchasing Power Parity
ProRural	National Strategy for Productive Rural Development (Nicaragua)
PRS	Poverty Reduction Strategy
PRSP	Poverty Reduction Strategy Paper
QSDS	Quantitative Service Delivery Survey
SISEVA	Evaluation System for Learning (Nicaragua)
SWAP	Sector-wide approach
WCA	World Programme for the Census of Agriculture

Preface

Having the capacity to track results and to use that knowledge to learn what does and what does not work – or how to make things work better – makes M&E a powerful tool for improving development processes and outcomes. In 2006, the Global Donor Platform for Rural Development (GDPRD) and the World Bank undertook to prepare this Sourcebook in collaboration with the Food and Agriculture Organization of the United Nations (FAO). The Sourcebook develops a framework for standardizing approaches for selecting indicators and proposes a menu of core indicators for monitoring and evaluating agriculture and rural development (ARD) activities. Ultimately, the objective is to improve the quality of monitoring and evaluation of agriculture and rural development programmes at the national and global levels.

M&E is intrinsically challenging and requires a level of technical capacity often unavailable in developing countries. The challenge is greater in the poorest countries and in post-conflict situations where less-than-optimal conditions, in particular, the weak statistical capacity, can cause major difficulties. This Sourcebook provides guidance on how to build the capacity needed for effective M&E in developing countries, starting with the identification and collection of the indicators. It suggests a number of approaches for determining which indicators to select given the different types of information that are most pertinent to different agricultural and rural activities, projects and programmes, and data availability. In addition, an innovative feature of the Sourcebook is the presentation of a core set of standard ARD indicators, with the recommendation that they should be regularly compiled by all countries. These "priority indicators" should be the same in all countries so as to allow for country comparisons, and to facilitate the monitoring of ARD programmes and goals at the international level. The Sourcebook identifies a core list of 19 priority indicators, as well as a menu of some 86 indicators that are categorized by sector, subsector and theme. It is hoped that countries may refer to and borrow from it when developing their own national ARD M&E programme. The menu of indicators was validated through in-country workshops in Cambodia, Nicaragua, Nigeria, Senegal and the United Republic of Tanzania.

This Sourcebook was prepared by a team of staff from the World Bank and FAO. Other member institutions of the GDPRD provided valuable inputs. Their remarks, as well as the analysis presented herein, will inform the ongoing GDPRD-facilitated dialogue among donors and partner governments on how to utilize statistics data to improve the management of agriculture, and to capitalize

on its special qualities as a high impact sector with regard to poverty reduction. The recommendations presented in this Sourcebook will also be applied in the *Code of Conduct for More Effective Agriculture and Rural Development Programmes* currently being developed by the GDPRD members.

The aid effectiveness agenda has put considerable pressure on all sectors to empirically demonstrate their performance. It is hoped that this Sourcebook will build upon practitioners' capacity to validate the effectiveness and impacts of agricultural and rural operations.

Christoph Kohlmeyer
Chair
Global Donor Platform for Rural Development

Juergen Voegele
Director
Agriculture and Rural Development Department
World Bank

Hafez Ghanem
Assistant Director General
Economic and Social Development Department
FAO

Executive summary

Background

At the United Nations Conference on Financing for Development, held in Monterrey, Mexico in 2002, both developing and developed countries made commitments to a shared responsibility to achieve development results, particularly those embodied in the Millennium Development Goals. Emphasizing results-based development requires the capacity to monitor indicators that reliably reflect results at all stages of the development process, from strategic planning to implementation to completion. Yet, donors and development practitioners still lack a common framework of results indicators to measure the effectiveness of development assistance. Developing a Monitoring and Evaluation (M&E) system that tracks these indicators using accurate and timely data is therefore a natural priority for the international development community as well as for developing countries themselves. For agencies and institutions involved in agriculture and rural development (ARD), this means developing a common framework that will enable donor agencies to harmonize their monitoring activities.

The reality is that many countries lack the capacity to produce and report the data necessary to inform the international development debate or to monitor their national trends. Although the situation is improving, global databases are still suffering from data gaps and inconsistencies as a result of weaknesses in National Statistical Systems (NSSs). In the final analysis, the validity of global monitoring systems depends on the quality of the data that comes from the countries. It is at the country level that problems occur, and it is at this level that assistance is required to build up sustainable capacity to collect and disseminate appropriate indicators.

Definition, objectives and methodology

Monitoring and evaluation are separate but closely connected activities. Monitoring is generally defined as a continuing activity that involves the collection of data on a regular, ongoing basis in order to track inputs, outputs, outcomes and impact while the project/programme is being executed. Evaluation, on the other hand, may use monitoring data, but is carried out at distinct and discreet moments of time to determine the worth or significance of a development activity, policy or programme. Taken together, they form a powerful instrument for planning the future on the basis of what can be shown to work and what does not.

Strengthening capacity for M&E at the subnational and national levels is intrinsically linked to M&E at the global level. Both depend on sound indicators

based on reliable and more complete data. To this end, the Global Donor Platform for Rural Development (GRPRD), the World Bank and the FAO set out to develop a menu of core indicators that could be used to monitor ARD at the project, national, regional and global levels. The approach is generic, but specific indicators are suggested that allow comparisons to be made **between** urban and rural areas, as well as **within** rural areas, specifically between agriculture- and non-agriculture-dependant communities and households. Separate sets of indicators are suggested for: the ARD sector as a whole; various subsectors (crops, livestock, forestry, fisheries and aquaculture, rural micro and small and medium-sized enterprise (SME) finance, research and extension, irrigation and drainage, agribusiness and market development); and related thematic areas (community-based rural development, natural resource management, and agricultural policies and institutions).

The purpose of this Sourcebook is to pull together into a single document a collection of common sense tips and recommendations based on actual practices and experience around the world. The Sourcebook aims first and foremost to help strengthen M&E capacity at the national and subnational levels, and to ensure a consistency of approach and methodology so that, at the global level, sufficient reliable and timely information can be accessed from the different countries and used to make cross-country comparisons and to calculate development indicators at the global level.

The ideal environment for establishing a good M&E system is where: (i) there is a strong and consistent demand for information; (ii) the concept of "management by results" is widely practised; (iii) timely and relevant information is systematically used to improve decision-making and to advance the process of development; and (iv) systems are in place to ensure that reliable and relevant information is available when needed. The less-than-ideal situation, on the other hand, is where (i) demand is weak; (ii) evidence is not used to inform decision-making; and (iii) the stock and flow of timely information are irregular and unreliable. The Sourcebook is specifically targeted towards countries where conditions are less-than-ideal, particularly with respect to the availability of relevant information.

Synthesis

The challenge of understanding reality on the basis of partial information is a recurring theme in the Sourcebook. It is particularly challenging in countries where conditions are less than ideal, that is, where the ability to collect and process statistical data is limited. The Sourcebook cautions against relying on a single source of information and encourages the use of the triangulation process – i.e. combining several sources of information to pick out the key elements of the story. In keeping with the theme of supporting M&E in less-than-ideal conditions, the focus throughout is on assembling recommendations that are pragmatic and practical, rather than abstract and academic. The Sourcebook emphasizes the need to keep things simple and suggests, for instance, that when countries assess their data

needs, they should focus on a minimum set of priority core indicators, rather than on a desired set. It looks at how indicators might be provided and used in conditions where data are limited and capacity to generate them is weak – a situation common to many countries. While the focus is primarily on the monitoring and evaluation of programmes in the agriculture and rural development (ARD) sectors, the guidelines are also relevant to other sectors. Indeed, the approach advocated in the document – which is strongly rooted in the idea of monitoring service delivery and measuring early outcomes – can be generally applied to almost all sectors, and provides an ideal basis for the monitoring of Poverty Reduction Strategies (PRSs) or other national development initiatives.

The Sourcebook reviews best M&E practices under three broad headings: the analytical framework, the data framework and the institutional framework.

Analytical framework

The analytical framework examines how one measures the impact of the development initiative. What indicators are needed and how are they selected? A complete M&E system must identify and monitor indicators at each of four levels – input, output, outcome and impact. Nowadays, most projects/programmes have a Management Information System (MIS) for tracking inputs and outputs (performance). A fundamental and essential output of the M&E system at this level should be the production of regular performance monitoring reports serving as an input into the preparation of annual work plans and budgets. Tools and approaches such as public expenditure tracking surveys are described in the Sourcebook.

Once systems are in place to monitor performance, attention can turn to the monitoring of results (outcomes and impact) – and this is the area on which the Sourcebook concentrates most. The shift in emphasis from performance to results has profound implications for M&E. Unlike performance monitoring, where data are relatively easily available from internal institutional information systems, measuring results involves turning to the targeted beneficiaries (clients) for information on the project and how it has affected them.

Changes in yield and production levels, whether for crops, fisheries, livestock or livestock products, inevitably feature among the main indicators used for monitoring project outcomes. The Sourcebook suggests that where objective measures are difficult to obtain at the early stage of interventions, farmers' own assessments can serve as useful proxies.

The Sourcebook also shows how a service delivery approach can be used to select indicators which can generate useful, easy-to-measure early outcome measures. It suggests that greater use be made of qualitative indicators, such as access, use and satisfaction.

Finally, there is the question of evaluation. This can be a seriously data-hungry exercise, but for countries with limited capacity, there are ways of getting around the problem. Not all projects/programmes need full-scale impact evaluations, and

where required, they may be carried out without collecting much additional data beyond what has been routinely collected for monitoring purposes – provided the evaluation is carefully planned in advance. Good evaluation will almost certainly involve combining data from various different sources and coming to a considered view on the impact of a particular intervention based on a triangulation process and weighing up of messages – often apparently inconsistent – from different sources.

Nevertheless, for most evaluations and broader planning purposes, the Sourcebook emphasizes the need for a set of basic agricultural and rural sector statistics that extends beyond the service delivery measures. These include basic sector statistics, such as area production and yield data, prices, agricultural input use, public spending on agriculture, the contribution made to GDP by agriculture and GDP per capita. In countries where these are not available, they should be put on a priority list for inclusion in any statistical capacity-building programme. An extended menu of indicators is supplied in Annex 1, which countries can use to help them prioritize and select the most useful indicators for their particular needs. The list is not exhaustive nor is it expected that all countries should adopt and use all the indicators, but it offers a choice and includes examples of good practices taken from different countries around the world.

The discussion of the analytical framework concludes with reference to monitoring and evaluation at the international level. It identifies a set of 19 priority indicators already included in the menu of indicators as core indicators for tracking ARD sector outcomes at the international level. These 19 indicators have been selected on grounds of comparability, availability and relevance. They represent a universal minimum core set and, as far as possible, should be included in all national M&E programmes. Without this minimal commitment at the country level, it is not possible to improve the quality of M&E at the international level, which is one of the purposes of the Sourcebook. But this should not be too onerous a burden, since the same indicators are used to monitor not only at the international level, but also at the national level.

Data framework

In order to meet the needs of monitoring at each of the four levels (inputs, outputs, outcomes and impact), the M&E system needs to draw on information coming from a variety of different sources. It is not just that each level requires different indicators, but also that the requirements of the users in terms of periodicity, coverage and accuracy vary according to the level of indicator. Input indicators are required to inform short-term decision-making. They therefore need to be produced frequently and regularly – possibly once every 1-6 months. The same applies to output indicators, but here the reporting period can likely be longer. As one moves further up the results chain and starts to collect more information about clients rather than the servicing institution, the task of data collection becomes more complicated. Time must be allowed for clients to become aware of and start using public services. One may see little evidence of outcomes

for the first few years. Therefore, it may be acceptable to build a programme around a reporting schedule of, for instance, 1-2 years. But it is important that the process is initiated at the very beginning of the project with a view to using the first report for establishing the baseline situation. The evaluation of the eventual impact comes much further down the line – often years after the project has been completed. Although the time frame may be more relaxed, the analytical challenge is not, and from the data collection perspective, experience teaches us that it is vital that the outline on how the project is to be evaluated is agreed from the very beginning, since it may involve setting up an experimental design to try to isolate the "with/without" project effect.

The Sourcebook devotes considerable attention to the need for a strong statistical infrastructure and reviews the range of different statistical instruments available.

The most popular and obvious instrument for monitoring outcomes of ARD programmes is the household survey. It provides data that can be disaggregated to show results for different population groups and has the advantage of providing information on both beneficiaries and non-beneficiaries. There are a number of different household survey models that can be used, each with its own strengths and weaknesses. The Sourcebook assesses their relative strengths and weaknesses and approximate costs. The most complete coverage is provided by the population census. Although obviously not appropriate for day-to-day monitoring, the census is important because it provides the framework for almost all other household survey activities, including agricultural censuses and surveys. The latter are extremely relevant to the monitoring of ARD programmes because they are usually the only means of monitoring changes in crop production levels and yields. Integrated multi-topic household surveys are another form of enquiry that has become increasingly popular. They are particularly good as baseline surveys that can be used to measure poverty levels, identify potential problems in need of attention, and generally understand the way in which households establish mechanisms to cope with difficult living conditions. The big disadvantage is that they are difficult surveys to undertake, and many countries have neither the analytical nor the survey capacity to successfully carry out such large-scale complex surveys on a regular basis. Lighter and more rapid household surveys are, however, becoming increasingly popular. Service delivery surveys have been used in market research for a long time, but are relatively recent additions to a National Statistical Office (NSO)'s repertoire of surveys. They are extremely well-suited to monitoring early results. They are also easy to implement and can be repeated annually without disturbing any other survey work that the NSO may be undertaking.

In addition to household surveys, a good M&E system will use a wide range of other tools. These can include community surveys, which may be conducted both on probability and non-probability samples, and qualitative surveys and studies, including participative assessments, focus group discussions and rapid appraisals such as windscreen surveys. Institution-based surveys, such as Quantitative Service

Delivery Surveys (QSDSs), can also play an important role in highlighting supply-side constraints, as can the analysis of administrative records.

The main message to emerge from the Sourcebook is that no single instrument can meet all needs and that any monitoring system will most likely acquire indicators from several sources – both formal and informal. Since it can take a while for the necessary capacity to be built, the Sourcebook offers a number of possible shortcuts for countries with less developed statistics systems.

In many countries, NSOs have found themselves caught in a vicious circle in which users have become disillusioned because the statistical products are late, inaccurate and filled with blanks. In a number of cases, this has led users to become dismissive of the efforts of the NSO, and in the process, to stop providing feedback on how databases could be improved. The inevitable knock-on effect is that resources for statistics are reduced and, as a result, so are NSO capacities. However, the future looks more promising and the signs are that with some assistance, NSOs will be able to rebuild capacity and meet the new information demands required by the monitoring of national development strategies.

Institutional framework

The final challenge in building up M&E competences is neither technical nor conceptual, but managerial. It concerns ensuring that the required incentive structure and institutional capacity are created to be able to perform this work. Whether countries already have an active ongoing M&E programme or whether they are starting from scratch, they need to regularly review all ongoing M&E activities. This may unearth a number of apparently duplicating and conflicting structures, but the goal should be one of inclusion not exclusion, and of creating a network of institutions engaged in M&E.

At the core, there needs to be a central M&E unit with the authority to coordinate the different initiatives. One of the more important functions of the unit should be to promote and encourage the demand for M&E. At the same time, it needs to help establish stronger links with data suppliers within the National Statistical System (NSS).

Despite the numerous areas of common interest, in many countries there appear to be two distinct and separate communities of practice – the M&E community and the statistics community. Both may be working on parallel issues but not necessarily communicating or working together. At the same time as the growth of interest in the M&E of national development programmes, there has been a similar interest in the rehabilitation of NSSs. The NSS comprises all the institutions and agencies that contribute in some way to the bank of national statistical data, which includes line ministries, Customs and Excise and the Central Bank, among others. The apex institution for the NSS is the NSO. Many countries are now developing National Statistical Development Strategies (NSDS) in such a way that they are integrated into national development policy processes. This ties in closely with the ideas underpinning the development of national M&E capacity.

The role of development partners

Donors have been among the strongest advocates for establishing good M&E procedures and for building up M&E capabilities. They have also provided strong support to the strengthening of national statistical capacity, but in many cases, their efforts have been counter-productive as a result of a failure in coordination. However, all major donors have now subscribed to the Marrakesh Action Plan for Statistics (OECD, 2004), in which donors commit themselves to working collaboratively to support countries in the preparation of NSDS.

Emerging issues

One cannot leave the discussion of the evolving role of M&E without making reference to three new and growing challenges. The first is the impact of devolution and decentralization on M&E. Many countries now pursue broad decentralization policies aimed at bringing the government closer to the people and enhancing transparency and accountability. This has profound consequences for M&E, which is now obliged to provide indicators at a much lower level of disaggregation. When the data source is administrative records, this may not present much of a problem. But when the source is a statistical survey, it can require dramatic increases in sample sizes, which may call for a major rethinking of how data are to be collected. The second challenge concerns the involvement of communities themselves in M&E. As interest in community-driven development projects continues to grow, so too does the demand for community-driven M&E in which the communities themselves take charge of their own M&E. This is likely to be an area in which major methodological developments will occur. Finally, there is the challenge of the monitoring and evaluation of ARD programmes at the global or international level. Monitoring international/global goals is the responsibility of the international development institutions, including the specialized agencies of the United Nations, the World Bank and the International Monetary Fund (IMF), but ultimately these entities depend on the NSSs to provide the basic data. The relationship between national and international institutions engaged in monitoring is not hierarchical, but rather, complex and symbiotic. Ultimately, the global M&E network is only as strong as its weakest link. International agencies therefore have a vested interest in seeing that the capacity of national institutions is strengthened.

Setting up an M&E strategy in agriculture and rural development

The Sourcebook makes the point that a fully evolved M&E system is more than a simple tracking system to measure performance and outcomes. These activities need to be put into the context of a cyclical approach to management in which:

- planning involves the articulation of strategic choices in light of past performance;

- implementation includes ongoing performance monitoring and periodic evaluation that provide opportunities for learning and adjustment;
- reporting on results is used both for internal management and for external accountability to stakeholders, including civil society. The reporting phase also provides managers and stakeholders with the opportunity to reflect on what has and what has not worked – a process of learning and adjusting that feeds into the next planning cycle.

The Sourcebook, in its final chapter, describes the key elements of an ARD M&E strategy and sets out the key steps that need to be followed to set it up, namely:

- Assessment of current M&E capacity and diagnosis.
- Review of indicators using the methodology described in Chapter 2 and, where appropriate, the suggested indicators provided in Annex 1.
- Review of current data, sources and gaps. The assessment should include a review of the quality and timeliness of the data and should draw on information contained in Chapter 3.
- Develop action plans linking together the M&E activities of all the institutions involved – as described in Chapter 4.
- Review resource requirements.
- Define a system to monitor the performance of the M&E action plan.

What is, in effect, being proposed in the Sourcebook is that countries should define a strategy for developing national M&E capacity as part of their overall ARD strategy. This would result in a better understanding of what works and what does not, which will lead directly to better planning of future programmes and projects. It will also lead to better programme implementation by providing timely warnings suggesting how resources may need to be reallocated when actual results are deviating from expected results.

Chapter 1
The evolution of M&E in development

The chapter opens with the question "What is M&E?" and then demonstrates how M&E has different meanings for different groups. The chapter then describes how M&E has evolved over the last 20 years from its early beginnings as a project-based evaluation tool to its current form, which is used for tracking multisectoral national development programmes such as Poverty Reduction Strategies.

What is M&E?

In the old story of the blind men and the elephant, a group of blind men touch an elephant to determine its true nature. Each one touches a different part. The one who feels a leg says the elephant is like a pillar; the one who feels the tail says the elephant is like a rope; the one who feels the trunk says the elephant is like a tree branch; the one who feels the ear says the elephant is like a fan; the one who feels the belly says the elephant is like a wall; and the one who feels the tusk says the elephant is like a solid pipe. They each claim to know what an elephant is but they are in complete disagreement. All are partially right, yet all are wrong.

The story of the blind men and the elephant could apply to M&E. Ask six people what M&E is and you get six different answers! It means different things to different people: M&E is a management tool; M&E improves planning; M&E is applied research; M&E is a tool to improve governance and accountability; it empowers communities; it monitors global goals. In fact, it covers all of the above and includes project supervision, financial monitoring, surveys and statistics, MISs, social analysis, and the setting and tracking of national development goals. Yet, it is more than the sum of its component parts.

The story of the blind men and the elephant is also relevant to M&E in another way. It illustrates how difficult it can be to understand reality on the basis of partial information. This underlines one of the key messages of the Sourcebook, which is to emphasize throughout the importance of sharing and triangulating information from different sources, and to be wary of relying on a single source of information. This applies equally to qualitative and quantitative information.

Different sources have their own individual strengths and weaknesses. In the area of poverty monitoring, for instance, the messages derived from qualitative studies based on participant observation often yield results that are seemingly at odds with the findings from "objective" statistical household surveys. The temptation is to reject one (usually the qualitative data) as being wrong. This would probably

Box 1. Definitions of monitoring and evaluation: 1984-2002

1984

Monitoring is a continuous assessment both of the functioning of the project activities in the context of implementation schedules and of the use of project inputs by targeted populations in the context of design expectations. It is an internal project activity, an essential part of good management practice, and therefore an integral part of day-to-day management.

Evaluation is a periodic assessment of the relevance, performance, efficiency and impact of the project in the context of its stated objectives. It usually involves comparisons requiring information from outside the project – in time, area or population.

IFAD, 2002

2002

Monitoring can be defined as "a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds". Thus, monitoring embodies the regular tracking of inputs, activities, outputs, outcomes and impacts of development activities at the project, programme, sector and national levels. This includes the monitoring of a country's progress against the Millennium Development Goals (MDGs) or other national measures of development success.

Evaluation can be defined as "the process of determining the worth or significance of a development activity, policy or program ….. to determine the relevance of objectives, the efficacy of design and implementation, the efficiency or resource use, and the sustainability of results. An evaluation should (enable) the incorporation of lessons learned into the decision-making process of both partner and donor."

OECD, 2002

be a mistake. The measurement and monitoring of living standards is a highly complex undertaking because of the multifaceted nature of the subject matter. When trying to interpret messages coming from different sources, it may at times seem as if one is trying to compare apples and oranges. Closer inspection and comparison of the two sources, however, often reveal important insights and show that far from contradicting each other, they actually highlight different aspects of poverty and provide complementary information. The key point is not to misuse any one instrument and expect it to answer questions that it was never designed to answer.

The first task of the Sourcebook, therefore, is to ensure that everyone has a common understanding of the issues that M&E can legitimately be expected to address. Various texts have defined M&E differently, which leads to more confusion. Among the earlier attempts, the clearest and least ambiguous definitions were found in the *Guiding principles of the design and use of monitoring and evaluation in rural development projects and programmes*, produced by IFAD in 1985 in cooperation with FAO and the World Bank (IFAD/FAO/WB, 1985). Box 1 compares the definitions established in 1984 with those revised and updated by the DAC Network on Development Evaluation (OECD, 2002) almost 20 years later.

The language is different but the concepts are broadly similar. What has changed, however, is the way in which the M&E concepts are applied. In the early days, the focus was on the project – a relatively well-contained development initiative with a limited time frame and clearly articulated goals. Today, however, the focus of M&E efforts is much broader and encompasses the M&E of sectoral plans and programmes, national development strategies, and, indeed, the international Millennium Development Goals.

Another important point to note is that, in both the earlier and the later definitions, the idea of M&E as an audit-like surveillance tool is excluded. Where there is an M&E unit, rather than being treated as an external agent, it is integrated into the project management structure and serves as a resource for supplying key information on project implementation and delivery. The function of the M&E unit is seen as assisting management by establishing and maintaining appropriate MISs and ensuring that they produce reliable data in a timely manner. Good management requires a good MIS and that the monitoring function is carried out using the data from within the MIS. Such a system includes the basic physical and financial records, the details of inputs and services provided to the beneficiaries or clients (for example, credit and extension advice) and data obtained from surveys and other recording mechanisms designed specifically to collect information from the service users.

M&E has evolved from being a set of project management tools to becoming a core element of national strategies for reducing poverty.

Evaluation is seen as a separate function but linked to monitoring. Evaluations can be simple or complex. There are several different kinds of evaluations, ranging from short desk reviews of documents and performance audits, to full-scale impact evaluation. Impact evaluation has a critical role to play in increasing knowledge about what works and what does not. Impact evaluations can be immensely valuable but are not easy to carry out. They draw on the MIS to provide data for making comparisons over time and against comparable "control" information, but they also require information from the clients – the intended beneficiaries. This requires baseline information. In the beginning, it was implicitly assumed that the project M&E units would undertake baseline surveys of their own with the understanding that the survey would be repeated at the end of the project and any differences would then be attributable to the project itself. In most cases, this proved to be much more difficult than anticipated. In many cases, the survey was overambitious and took years rather than months to complete. At times, the second survey was never undertaken, or if it was, the size of the combined sampling and non-sampling errors was found to be larger than the real change that the surveys were meant to detect.

Monitoring and evaluation are closely linked but separate activities.

Even today, the relationship between monitoring and evaluation continues to be the subject of discussion. At one end of the spectrum, there are those who put the primary focus on monitoring, and see M&E principally as a management support system whose main concern is to ensure the timely production of appropriate indicators. At the other end of the spectrum, there are those with an equally strong argument that the primary function should be to carry out effective impact evaluation from which lessons could be learned for the future. Then there are those who feel that M&E systems should be capable of doing both. This middle path is the one that is usually taken – a sensible compromise where one must, however, be continually aware of the risk of spreading resources too thinly in trying to achieve multiple objectives and ultimately satisfying none.

To summarize, the basic principle is that monitoring is an ongoing activity and evaluation is periodic, carried out at specific times during the project cycle (annual, mid-term, terminal) or indeed after the completion of the project (impact evaluation). In broad terms, M&E are activities whose primary function is to provide appropriate information at the right time to users with decisions to make and to improve their decision-making as a result. M&E, like all other services, can only function effectively if there is a demand for

The measure of a good M&E system is customer satisfaction.

it. How can one know whether the system is working correctly or not? In the long term, one would seek evidence of better planning, resource allocation and administration of development programmes as a result of learning from experience. In the short run, the answer is satisfied users. If there is a growing number of people who are aware of M&E data and also a growing number of people actually using the data, then one may infer that the system is providing a useful service.

Who are the users?

The more open or inclusive the system of government, the broader the range of users is likely to be. At the start, the focus of the M&E reporting system may be on budget management and performance budgeting, but as the programme or project grows and the number of beneficiaries increases, so does interest in the M&E data. Users include those who have a financial or management interest in the project (donors, government), as well as the beneficiaries, the media, civil society at large and their representatives (parliament).

At the beginning, however, it can be hard to raise any interest at all. In the early days, in many countries, the demand for good M&E information originated entirely from outside sources. The donors were driven by an electorate at home that needed to be satisfied that aid funds were being used for the intended purposes and were achieving results. In the developing countries in which the M&E systems were being installed, however, there was generally little interest. Even in projects that included a donor-driven M&E component, managers were ambivalent about its value and tended to see M&E units as a drain on their resources, or even worse, as an informant imposed from outside. We have moved a long way since then, but still, without in-country demand, no system can be sustainable. Therefore, one of the first requirements for successful M&E is to nurture and cultivate the demand. This is likely to mean taking measures to initiate a strong advocacy programme to inform potential user groups about the benefits of a results-driven environment. Consequently, M&E has become an important pillar of the PRS and not just a marginal activity. As shown in Box 2, the PRS can underline the need for good M&E data to: (i) support budget decision-making; (ii) help with policy formulation and programme development; (iii) support the management of sectoral programmes; and (iv) signal whether the programmes are genuinely contributing to an improvement of living standards and well-being in the country. However, the process of reorienting a country or culture to value a results-oriented government system can be a long and arduous process. In summary, monitoring information and evaluation findings can contribute to sound governance in a number of ways, but primarily through evidence-based policy-making (including budget decision-making), policy development, management and accountability.

Box 2. How M&E findings help governments and stakeholders?

M&E findings:

- support policy-making, especially budget decision-making, performance budgeting and national planning. These processes focus on government priorities among competing demands from citizens and groups in society. M&E information can support the government's deliberations by providing evidence of the most cost-effective types of government activity, such as different types of employment programmes, health interventions, or conditional cash transfer payments. Terms that describe the use of M&E information in this manner include evidence-based policy-making, results-based budgeting and performance-informed budgeting;
- help government ministries in their policy development and policy analysis work, and in programme development.;
- help government ministries and agencies manage activities at the sector, programme and project levels. This includes government service delivery and staff management. M&E identifies the most efficient use of available resources and can be used, for example, to identify implementation difficulties. Performance indicators can be used to make cost and performance comparisons – performance benchmarking – among different administrative units, regions and districts. Comparisons can also be made over time that help identify good, bad and promising practices, which can prompt a search for the reasons for this performance. Evaluations or reviews are used to identify these reasons. This is the learning function of M&E and is often termed "results-based" or "results-oriented management";
- enhance transparency and support accountability relationships by revealing the extent to which the government has attained its desired objectives. M&E provides the essential evidence necessary to underpin strong accountability relationships, such as the government to the Parliament or Congress, civil society and donors. M&E also supports the accountability relationships within government, such as between sector ministries and central ministries, among agencies and sector ministries, and among ministers, managers and staff. Strong accountability, in turn, can provide the incentives necessary to improve performance.

World Bank, 2007

How M&E has evolved

At this stage, a historical learning exercise may be useful. In the following description of how M&E has evolved over recent decades, this process has been grouped into several distinct phases for the purpose of clarity. This is an oversimplification and disguises the fact that progress is neither sequential nor linear, but it does help to show how ideas have evolved and how expectations have expanded over the years.

In the beginning: project-based M&E

The first signs of interest in M&E for ARD projects became evident in the mid-1970s. At that time, interest was strictly project-based and there was general agreement that projects could be better designed and managed with a strong M&E programme. In many cases, this involved the establishment of a dedicated M&E unit.

In the 1970s, interest in M&E was strictly project-based: its main purpose was to serve as a management tool.

During this early evolutionary phase of M&E, its main purpose was to serve as a management tool that would provide timely feedback and give warning whether the project was on track or not. While paying lip service to the need for measuring outcomes, the focus of interest was on the monitoring of inputs and outputs. The project document was treated more like a "blueprint" than a "roadmap". If the planning had been correctly done, then the main purpose of M&E was to provide timely feedback that the project was being implemented in line with expectations – and if not, to send a quick warning. This is still an important aspect of M&E even today.

Expanding horizons: programme and sectoral M&E

By the early 1990s, a change was taking place in how development aid was being administered, leading to a shift in focus from the project to the sector-wide programme. Programmatic aid, whether in the form of loans or grants, was becoming increasingly common, since it was seen that project-based assistance was failing to deal with the larger systemic problems and was not creating an effective investment-friendly environment necessary for sustainable development and long-term raising of living standards. The effect was not so much that projects were discontinued – indeed they continued to thrive – but that a sector-wide approach (SWAP) became increasingly popular as a means of promoting and coordinating sector-wide and national development planning. These development models potentially gave more flexibility to governments and programme executing agencies, but good reporting and feedback systems had to be conceived of as an integral part of the programmes.

The expansion from project to programme-level support had enormous implications for the M&E system.

One of the results of this SWAP was the recentering of many M&E activities from the project level to the sectoral level. Monitoring and evaluation became functions of sectoral ministries and appropriate M&E units were established at the ministry level. Sometimes, the sectoral units entirely replaced the project units; sometimes they did not. A network of M&E units were created, in which project units either copied their reports or sent them directly to the sectoral M&E unit. The nature of the relationship between the project units and the sector unit varied substantially from country to country. In some, it was rigid and hierarchical; in others, the relationship was much looser. But the old custom of allowing each project to design and develop its own M&E procedures was in general replaced with a more centralized approach that would ensure that all programmes and projects followed the same procedures and reporting formats so that statistics could be compiled into sector-wide reports. Development partners also had to be prepared to accept a standard format rather than insist that their own individual reporting formats be used.

The focus turns to the beneficiaries, which requires better data and more tools.

In the 1990s, the idea of results-based management was also becoming popular. The consequence was a shift in emphasis away from the monitoring of inputs and outputs to the measurement of "results" – a much more difficult task. This expansion of expectations was a significant change from before. Up to that time, it was possible for much of the data to be generated from internal reporting systems. Then, in order to measure the results of project activities, the focus of M&E had to switch from the project to the client or intended beneficiary. It thus became necessary to call on a much wider range of data tools and sources. Surveys and beneficiary interviews in particular would need to be undertaken, which required a level of expertise and training not generally available in project M&E units, or even in the M&E units of sector ministries. For the most part, M&E staff did not have the time, training or the resources to tackle this kind of work. The involvement of new players with more technical expertise was needed.

The early involvement of NSOs was not particularly successful.

One new player was the NSO. The primary function of an NSO had always been to act as the ultimate source and repository of all official national statistics. In most countries, they were established as a government body with only limited autonomy. Their most important outputs were national accounts, an annual statistical abstract and the published results of whatever survey or census they happened to have undertaken recently. In many countries, it seemed to be the only institution with the knowledge and capacity to collect and process data on the scale needed by the project. It was thought that either it would be possible for projects to "piggyback" onto the NSO's household survey infrastructure

and to use the NSO survey as a means of measuring project results, or it could undertake special surveys specifically for the project. In both cases, the outcome was generally disappointing. Statistics offices were, on the whole, overextended and under-resourced, and failed to rise to the challenge. Adherence to timeliness and respect for deadlines were not qualities commonly associated with under-resourced NSOs – nor was adaptability. Another problem was that the data supplied were generally too "macro" and not sufficiently disaggregated for M&E purposes. While their data could make a contribution to the overall performance of national and sectoral development programmes, they were generally not specific enough to be helpful in measuring the outcome of specific development interventions. Either they were not repeated with sufficient regularity to allow for comparisons over time, or they could not be sufficiently disaggregated to allow for comparison between different subgroups of the population. The dialogue between the national data provider and the data user was not easy, and led to frequent disappointments.

The arrival of poverty monitoring

Another force that started to emerge during the mid-1990s was concern about the issue of poverty. While the primary goal for a developing country had traditionally been "development through growth", it changed in the 1990s to "growth and poverty reduction"; it was not enough to aim for wealth alone. It now became increasing clear that this had to include a fight against poverty and protection for the poorest. A new branch of monitoring activity was required: poverty monitoring. This was a complex and challenging undertaking that, for the most part, was built around the tracking of living standards with a view to anticipating the direction in which they would likely move as a result of macroeconomic policy. One of the underlying driving forces was the concern that structural adjustment programmes – introduced in many countries in order to redress economic imbalances and improve international competitiveness – could be imposing undue hardship on some of the most vulnerable elements of the population.

In the 1990s poverty monitoring was introduced to study the effect of economic development on living standards.

Most countries started out with very little knowledge or capacity to monitor poverty. Support from donors focused primarily on assistance in the design and implementation of multi-topic household surveys, which included the measurement of household consumption as the indicator of choice for measuring poverty. Only NSOs had the capacity to undertake such large-scale national household surveys, but even then, in most cases, they did not have the capacity to analyse them.

As time progressed, qualitative and quantitative tools were added, including participatory poverty assessments, poverty mapping and the tracking of core

indicators over time. These involved bringing on board other institutions, including academic institutions and NGOs. To coordinate all these activities, countries started to establish National Poverty Monitoring Units. The results of these efforts were mixed, but overall capacity was being built. What is interesting, however, is that the building up of a national poverty-monitoring capacity was kept distinct and separate from other M&E capacity-building efforts, and there was very little communication between them … until the new millennium.

Monitoring Poverty Reduction Strategies: building national M&E capacity

By the turn of the millennium, poverty alleviation had moved from being a marginal issue to being a central concern for almost all countries. A target of halving global poverty by 2015 was enshrined as the first Millennium Development Goal. At the country level, the National Poverty Reduction Strategy (PRS) was introduced to serve as a framework for promoting the vision of "pro-poor growth" (Box 3). The earlier experiences of setting up country-level poverty monitoring systems were to prove critically important for the introduction and successful implementation of national PRSs. The poverty assessments provided the means of identifying where the most vulnerable were located.

M&E becomes a key agent of development in its own right.

The new millennium saw the bringing together of project- and sector-based M&E efforts with poverty monitoring activities. The result was the emergence of national M&E programmes centered around the monitoring of PRS results. At this stage, M&E started to emerge as a key agent of development in its own right, and an essential component of the PRS. In-country demand, which had previously been limited, started to expand – and with it, recognition emerged that M&E information should be not just a tool for policy-makers and planners, but should be made readily available to members of the public and to civil society. In this way, the M&E system started to become a tool for promoting good governance and accountability.

Coordinating M&E activities across and within sectors remains a challenge.

Box 3. Poverty Reduction Strategy Papers (PRSPs)

Poverty Reduction Strategy Papers (PRSPs) are prepared by governments in low-income countries through a participatory process involving domestic stakeholders and external development partners, including the International Monetary Fund (IMF) and the World Bank. A PRSP describes the macroeconomic, structural and social policies and programmes that a country will pursue over several years to promote broad-based growth and reduce poverty, as well as external financing needs and the associated sources of financing.

What is the purpose of PRSPs?

The world economy has grown steadily in recent decades, bringing widespread prosperity and lifting many millions out of poverty, especially in Asia. Nevertheless, in the next 25 years, the world's population is projected to grow by about two billion people, most of whom will be born in developing and emerging market economies. Without concerted efforts by countries to help themselves through sound policies and by the development community to increase its support of the countries' own efforts, many of these people will be doomed to poverty.

The PRSP approach, initiated by the IMF and the World Bank in 1999, results in a comprehensive country-based strategy for poverty reduction. It aims to provide the crucial link between national public actions, donor support and the development outcomes needed to meet the United Nations' Millennium Development Goals (MDGs), which are aimed at halving poverty between 1990 and 2015. PRSPs provide the operational basis for Fund and Bank concessional lending and debt relief under the Heavily Indebted Poor Countries (HIPC) Initiative. They are made available on the Web sites of the IMF and World Bank by agreement with the member country.

Core principles of the PRSP approach

Five core principles underlie the PRSP approach. Poverty reduction strategies should be:

- **country-driven**, promoting national ownership of strategies through broad-based participation of civil society;
- **result-oriented** and focused on outcomes that will benefit the poor;
- **comprehensive** in recognizing the multi-dimensional nature of poverty;
- **partnership-oriented**, involving coordinated participation of development partners; (government, domestic stakeholders, and external donors);
- based on a **long-term perspective** for poverty reduction.

IMF Factsheet, September 2005

Chapter 2
The analytical framework

This chapter deals with the classification and selection of indicators. The logframe is used to differentiate between project inputs, outputs, outcomes and impact. Indicators are needed at each level for effective monitoring and evaluation; each have their own defining characteristics and are discussed in turn. Tools to facilitate the collection and use of such indicators are reviewed. The main focus of the chapter is, however, devoted to outcome and impact indicators, and to the measurement of results, in particular, early results. The Sourcebook suggests that, for early results, a service delivery approach can work well. For longer-term results and impact measurement, a menu of core statistics is proposed. The chapter concludes with recommendations for selecting indicators for the ARD sector as a whole and for the various subsector programmes. Nineteen priority indicators are proposed. The process may also be assisted by reference to Annex 1, which contains a menu of potentially useful indicators.

Thinking logically about indicators

A good M&E system should, in principle, be integrated into all stages of a project or programme cycle, from identification through the evaluation. At each stage, it should seek to answer the question, "Are we on track?" At the end, it should answer the question, "Did we achieve what we wanted to achieve?" Throughout the duration of the project, the M&E system should generate timely reports on project progress, sounding alarms where necessary, and providing project management with the necessary information to help keep the project running as smoothly as possible. In the end, sufficient information should have been accumulated for an evaluation to be conducted to inform the appropriate stakeholders on whether the project had achieved its expected objectives and to highlight any unexpected outcomes. This is what should happen – in principle.

A project or strategy preparation team will find the situation on the ground much more complex. Development is the result of a complex interaction of forces that cannot be easily summarized as a simple flow of causes and effects. Most development goals are achieved as the result of a number of different interacting interventions. Much of the M&E literature places a heavy emphasis on the "evaluation" aspect of M&E. It suggests that the purpose of M&E should be to measure the extent to which the development goal has been achieved and then identify the contribution made by each intervention or project. In practice, just getting an answer to the question – "Are we moving in the right direction?" – is difficult enough. Answering the question – "Are there better ways we could be moving?" – is almost impossible. In the real world, the problem is that, in most cases, the data are just not available to carry out the kind of analysis that in principle seems so logical.

When choosing indicators, the starting point should be the question, "Is this proposed indicator measurable?"

A great deal has been written on the selection of appropriate indicators, and extensive lists have been prepared suggesting suitable indicators for monitoring different types of projects. These are useful reference materials, but in many cases, impractical to apply. Not only are there hundreds of indicators, but also the data that underpin them usually cannot be secured with the necessary precision or regularity. When choosing indicators, the starting point should be the question, "Is this proposed indicator measurable?" This helps considerably in the quest to identify a minimum list that requires the lightest of M&E structures. Even so, the range of possible indicators is still sizeable, which reflects the fact that the M&E systems still have to satisfy the needs of a broad range of users, which are not identical by any means. Annex 1 is there to serve as a checklist – a menu offering a selection of indicators. The actual selection of indicators should be a reflective and participative activity involving the key stakeholders who are most intimately associated with the project design and implementation – not an imposition of demands from outside. This chapter outlines a systematic approach that can be adopted to help prioritize the most critical indicators that need to be selected. It provides examples of how the methodology can be applied and used for different ARD subsector programmes.

A systematic approach can help prioritize the selection of the most critical indicators.

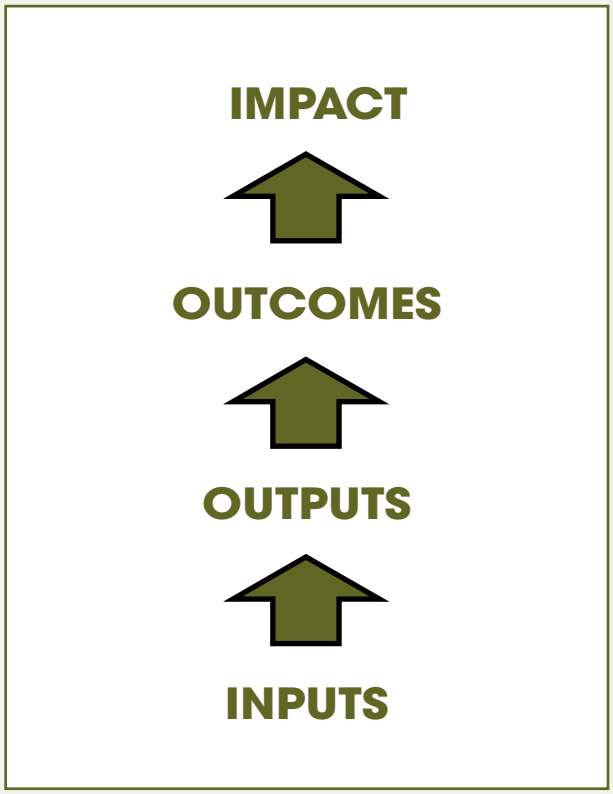
But first a word of caution. The number of indicators and the data required to compute them can grow rapidly. Even though there will always be good reasons for which the list of indicators needs to be expanded, there are also good reasons for starting small and making use of whatever data are available before collecting more. The Sourcebook strongly encourages the idea of integrating statistical

capacity building into national M&E programmes from the beginning, so as to ensure a reliable supply of core statistics from which the required indicators can be extracted.

The focus of this chapter is on indicators, but indicators are only signals. They can be helpful in highlighting whether the project or programme appears to be moving (or to have moved) in a particular direction, but they are, at best, rough instruments that can easily give wrong impressions and lead to misdiagnosis. Indicators alone are not sufficient for serious evaluation. They are merely the first step in a potentially complex and time-consuming analytical exploration. Good M&E also involves blending qualitative and quantitative information that together can enhance understanding of the situation on the ground.

Indicators are still only rough instruments.

The methodology for selecting indicators is initially introduced in the context of a project-level M&E system, but the process is the same even if one is working on indicators for monitoring a national PRS. The starting point is to establish a framework using the widely used logical framework approach (*logframe*). In very simplified terms, this is a conceptual device that describes the project in terms of its intended goal or *impact*. In order to achieve this impact, people's behaviour is expected to have changed in a way that will help with the achievement of the project goals. These behavioural changes are known as the project *outcomes*, and it may take several years before they become apparent. In order for these outcomes to occur, the project must generate *outputs* (goods and services). These outputs in turn require that the necessary combination of *inputs* (financial, physical and human) become available at the right time, place and quantity. Thus, in reverse order, the inputs will generate outputs, which will yield outcomes and eventually an impact. Take for instance the example of a small-scale irrigation project. Inputs in the form of staff training, equipment, and capital are used to generate outputs in the form of irrigation infrastructure, establishment of extension service, farmer training courses and research on improved crop varieties. The outputs then have to be made accessible to, and used by, the farmers whose changed farming practices in turn will generate outcomes in the



form of improved yields. Finally, these outcomes should lead to a positive impact in the form of higher revenues and greater food security.

The logframe is well known as a tool for project design and is a useful aid to better understand the logic that defines the development process. It has a second application, however, which is to provide the framework for developing a project M&E system that includes all stages of the project from beginning to completion and beyond. Once the logic of the project had been defined using the logframe, it should then, in principle, be a relatively simple process to monitor progress at each of the four levels. This idea has immense appeal because it helps to reduce the information needs for monitoring the project's success down to a relatively small number of key indicators, which, as already noted, is a desirable feature.

Logframe is useful and effective tool but has limitations.

The logframe does have its limitations, however. First, it promotes a blueprint approach to development. Project design can become a relatively inflexible and uncreative activity. Second, it reduces the process of development to a two-dimensional cause-and-effect formula – clearly a gross simplification. The third is that the project is conceived as an isolated entity and the complex interactions between projects with complimentary or competing goals tend not to be recognized, nor is the relationship between the project goals and the country development goals.

Nevertheless, the logframe can be effective, as evidenced by the fact that it has been widely used for a number of years and has heavily influenced the design of M&E systems. These systems have been most effective at the lower end of the causal chain, in monitoring inputs and outputs. As the project progresses, however, the functions of the M&E system change. This link to the project cycle provides a very useful framework for deciding what information is needed, when and for what purposes.

At this point, it will be useful to introduce two further concepts: **performance** and **results**. These are terms that were introduced after the logframe had popularized the notion of inputs, outputs, outcomes and impacts.

Performance refers to implementation or efficiency, and measures *actual* against *expected* results; it is a proxy measure of the quality of management. In general, it covers all four levels of the logframe causal chain, but focuses mostly on the bottom-end inputs and outputs and on how efficiently the project can convert inputs into outputs. Sometimes, the concept of performance is extended to include outcomes as well.

Results are the outputs, outcomes or impact of a development intervention. Results include the effects the project goods and services have on targeted beneficiaries and others. They may also include the negative effects, such as on the environment. Results are generally, but not necessarily, longer term and more complicated to measure than performance indicators.

Initially, the focus of M&E systems was on monitoring performance (i.e. a concentration on the lower-level input/output indicators), but with the growth of interest in "results-based development", it shifted to a higher level towards the monitoring of outcomes and impacts. A complete project M&E system should include the monitoring of both performance and results.

Monitoring performance (inputs and outputs)

Tracking inputs and outputs

The monitoring of project performance is M&E at its most basic level. It is the tracking of human, physical and financial resources and the recording of how they are converted into outputs (project goods and services). Strictly speaking, it includes financial monitoring and the analysis of financial records. In addition to generating financial reports, the data are used for cost-benefit analysis and analysis of costs per unit of output, etc. Cost data also lend themselves fairly easily to aggregation and merging with other data sets at higher levels. It is therefore relatively straightforward to integrate performance monitoring indicators into higher level (regional or global) tracking systems. Input and output indicators are generally simple to construct, and most of the information is readily available in project accounts and records. These are usually stored and disseminated through a Management Information System (MIS) that may or may not be connected to the financial management system. Information stored in the MIS includes data on unit costs (costs per hectare or per kilometre, etc.) and can also be useful for analysing the links between inputs and outputs, calculating key input/output ratios and for monitoring projects/programme performance and efficiency. The key to successful operations of the MIS is the ease with which data and monitoring indicators can be accessed and used by project management and others.

Performance monitoring is an essential part of good management.

Regular M&E reports should be generated at least annually and timed so as to serve as an input into the preparation of an Annual Work Plan and Budget. The allocation of budget resources of the following year should, in normal circumstances, be heavily influenced by the results and performance of the project during the current year – as recorded by the M&E system. Performance monitoring is now well established, particularly in projects receiving significant external funding.

A fundamental output of the M&E system at this level should be the production of regular performance monitoring reports.

Tools for monitoring inputs and outputs

At its most basic level, performance monitoring (inputs and outputs) is essentially a matter of "keeping the books". Proper and systematically maintained financial records are the starting point. At one time, they used to be maintained by hand, but are now handled electronically using an appropriate commercial financial management package.

Financial and management information systems

For most development projects that receive external financial assistance, it is perfectly satisfactory, indeed recommended, to use an off-the-shelf package as long as it can handle multiple currencies. In the early days, projects were given *carte blanche* to use whatever package they preferred. In an effort to improve the standardization of procedures, a number of countries now specify that public service institutions all use a single, nationally approved package. In addition to bookkeeping, the more general task of reporting on activities and outputs is required. But again, at its simplest level, this involves the establishment of simple reporting procedures and the collation of results into progress reports. As with the accounts, this could be done manually, but is now largely handled on the computer using an MIS. The choice of which system to use is a little more complicated, since it depends more on the nature of the project/programme. In general, the tools needed to operate the basic performance monitoring system at the project level need not be too complicated, and may even become easier as further technical advances are made.

Effective monitoring, open reporting and transparency strengthen local government and support the devolution of responsibility to local authorities.

Integrated local government information systems

When it comes to tracking sector- and subsector-level inputs and outputs, one finds significant variations from one country to the next, but the trend is shifting from a largely uncoordinated and disparate collection of project and sector monitoring systems towards the installation of a single coordinated set of procedures. This process has been assisted by the dramatic improvements in "connectivity" technology. Coupled with improved connectivity is the need to have a well-designed MIS that is adopted universally by all government offices, both at the national and subnational levels.

The United Republic of Tanzania is a country where such a programme is being successfully implemented under its Local Government Reform Programme (LGRP). The aim of the LGRP is to strengthen delivery of public services at the local level by a process of devolving administrative responsibilities to the local government authorities (LGAs) and making them the main conduit through which nearly all government and public services are channelled to rural areas.

Box 4. The national management information system for local government reform of the United Republic of Tanzania

Tanzania's local government reform programme (LGRP) aims to strengthen local authorities and transform them into effective instruments of social and economic development at the local level. It aims to improve quality, access and equitable delivery of public services, particularly to the poor, and thereby contribute to the government's efforts of reducing the proportion of Tanzanians living in poverty.

A critical component of the programme is the adoption of information and communications technology (ICT) and the development of a management information system (MIS) to facilitate the dissemination of reliable, accurate and timely information to a number of stakeholders, both within and beyond the government system. The MIS contains a number of separate systems, two of the most important of which are the Planning and Reporting database (PLANREP) and the Local Government Monitoring Database (LGMD).

PLANREP enables all local authorities to:

- create a performance budget framework of objectives, targets and activities;
- link any target to the national strategy for growth and poverty reduction (MUKUKUTA) cluster strategy;
- calculate projected revenue from formula-based and other grants from central government, own sources, the community and development partners;
- allocate conditional projected revenue to performance budget targets;
- allocate unconditional projected revenue to local authority departments and sections;
- export budget information to the Ministry of Finance;
- enter expenditure from manual or electronic accounting system;
- enter reports on the physical implementation of development targets.

LGMD is a local government monitoring system for capturing and reporting service delivery and socio-economic profile data. These data include information on education, health, agriculture, lands and water. It is also used to capture data from villages, wards and districts. The data are used to calculate 90 indicators. Data from the local authorities are forwarded to both the region and the centre for aggregation. These tools are being introduced to all local government authorities, albeit in a phased approach depending on the issues of local capacity, ongoing support and development of the systems. The software systems, infrastructure and equipment is simple to use and robust, and has been a good support system.

A key element of the LGRP is the development of MISs and the information and communication technology (ICT) infrastructure for the LGAs. Another key feature of the MIS is the development and support of systems that allow LGAs to collect, process and use the data needed for their own purposes and other local government stakeholders (Box 4).

When complete, the LGRP will make it possible for all districts to use the MIS to develop their own plans; prepare their own budgets; review their budget allocations; track expenditures; monitor their outputs in terms of the quantity of goods and services provided; and produce regular quarterly and annual reports – all with the help of the MIS. The country vision is for effective monitoring, open reporting and transparency that will contribute to more effective implementation of national strategic plans and improved governance.

Public Expenditure Tracking Surveys (PETSs) and Quantitative Service Delivery Surveys (QSDSs)

Not all countries are as advanced in the establishment of their M&E infrastructure as the United Republic of Tanzania, however; other solutions must therefore be sought under the less-than-ideal conditions where financial accounting systems are not functioning well. In such cases, countries have been undertaking Public Expenditure Tracking Surveys (PETSs) to track the flow of public funds and determine the extent to which resources actually reach the target groups. PETSs examine the manner, quantity and timing of releases of resources to different levels of government, particularly to the units responsible for the delivery of social services such as health and education. While a PETS traces money through the organization, a Quantitative Service Delivery Survey (QSDS) works to identify organizational weaknesses that can be addressed through reform. QSDSs address the issue of service delivery from the perspective of the supplier. These are surveys based on a random sample of facilities or service providers that focus on quality of service, characteristics of the facilities, their management and incentives structures. One output of the survey instruments is a case-by-case diagnosis of public service delivery, helping to identify weaknesses in implementation capacity and suggesting where reform efforts should be concentrated. PETSs and QSDSs are useful for diagnosing problems in service delivery and for providing evidence on delays, "leakage" and corruption in situations where little financial information is available.

Possibly, the most basic performance monitoring activity for sector-level programmes is the tracking of public expenditure.

Measuring results (outcomes and impact)

This chapter now shifts from performance monitoring to results measurement, now concentrating on higher-level indicators. It is at this level that the

demand for core indicators is strongest. A results-based system attaches the highest importance to providing feedback on outcomes and goals, rather than on inputs and outputs. In fact, with the advent of results-based management, there has also been a subtle but significant change in terminology whereby the terms "outcomes" and "impact" are frequently replaced by "early results" and "long-term results". The difference is slight, although the more recent terms better capture the time dimension. Both are used interchangeably in this Sourcebook. Box 5 presents the chief characteristics of the

Measuring results means turning the spotlight on the intended beneficiaries.

Box 5. Characteristics of different classes of indicators

	PERFORMANCE (Efficiency of the project or programme)		**RESULTS (Changes resulting from the project or programme)**	
LOGFRAME LEVELS	INPUTS	OUTPUTS	OUTCOMES	IMPACT
M&E ACTIVITY	Monitor resources and activities.	Track delivery of goods & services.	Assess early results (access, use and satisfaction with respect to services by users).	Evaluate long-term results.
CHARACTERISTICS OF INDICATOR	These indicators relate to physical, human and financial resources. Sources are MIS and administrative records.	Outputs are generated by the project/ programme. Outputs may include physical outputs, services, training, advice, etc. Sources include the MIS and administrative records.	Indicators should respond quickly and be easy to measure. They should measure the extent to which beneficiaries have changed behaviour due to project. Typical indicators include access, use and satisfaction with respect to project services. Sources include surveys of beneficiaries and service providers and service delivery data from surveys and administrative records.	Indicators may move slowly and be difficult to measure. They must show evidence of change and analysis must establish the extent to which change is attributable to project/programme being evaluated. They are derived from ongoing monitoring activities plus dedicated evaluation studies.
FREQUENCY OF REPORTING	Quarterly to annual.	6-18 months.	1-5 years.	5 years and over.

different classes of indicators and shows how the "results" terms fit with the more traditional logframe terms.

The shift in emphasis from performance monitoring to results monitoring has profound implications for M&E. Unlike performance monitoring, where the data are relatively easily available from internal institutional information systems, results monitoring turns to the targeted beneficiaries (clients) for information on the project and how it has affected them. A key objective of monitoring outcomes (results) is to highlight who is benefiting from the development programme or intervention, and how. At the same time, it is also important to know about the clients who are not benefiting and to understand why. This needs to be done while the programme is being implemented so that corrective action can be taken – simple in principle, but not so easy in practice. To make the task easier, it has now become good practice to separate the monitoring of short-term (or early) indicators from the monitoring of medium- to long-term indicators (which equate more closely to indicators that would be used to measure impacts). For the early indicators, rapid reporting now becomes a critical factor, which in turn affects the choice of indicator. Indicators that change slowly are not good indicators for measuring short-term outcomes, nor are those that are subject to extreme random fluctuations, that exhibit a long time lag or that take time and are expensive to measure. What are needed are indicators that respond quickly and that are easy to collect. Again, they should all be able to be disaggregated and presented for different subgroups of the population (e.g. by gender, vulnerable population groups, or the poor) and also be aggregated upwards and used to calculate indicators at the national, regional or global level.

Early results/outcomes

What, then, are examples of good indicators of short-term results? An examination of recent World Bank ARD Project Appraisal Documents (PADs) showed that project preparation teams have serious problems in identifying suitable indicators. There is a tendency to jump straight from performance monitoring to long-term outcomes. This leaves an important gap in the logical chain, which has sometimes been referred to as the "missing middle". The problem is that there is a time lag between the provision of project outputs and the outcomes on the target population; the result will not be felt in time to take corrective action – often not until several years after the project is complete. Such indicators are therefore of little value for providing quick feedback on early results: they either move too slowly or, due to their complexity or cost, can only be collected every five years or so. In the long run, it is clearly essential to have some objective quantifiable measure of the project impact – for instance, an increase in agricultural and non-agricultural rural

Monitoring service delivery is the key to tracking early outcomes.

income – but some other measure is needed in the short run, as it is impractical to think that such information can be collected and supplied on an annual basis.

So what can be done to fill the gap and catch the early signals of change? What sort of indicator can one use to measure short-term results? How can we know who have benefited from the project or programme and who have not? One solution is to ask the clients directly to evaluate how useful they feel the programme services have been. Consumer satisfaction is, after all, the standard measure used in market research to improve the quality of service delivery. So why not use a service delivery approach for monitoring development activities?

Access, use and satisfaction

A service delivery approach considers that most projects have one thing in common: they are essentially vehicles for making a product or products available to a target population. The concept of the "product" is a broad one, which may include:

- a tangible product such as a loan, a rural road, or a package of technological innovations for increasing yields;
- a service, such as an extension programme, local health care, or land registry service;
- something more abstract, such as "an enabling environment" or a "community development project".

It may even be a combination of the above – a package of products and services that the beneficiary might be expected to adopt. Even policy reform programmes can, with a little adjustment, be viewed through the service delivery lens. For instance, a decentralization policy should result in improved public services to the rural areas. These services are essentially the "product" resulting from the policy.

At its most simple level, a project comprises two elements: a **product** and a **delivery system**. For the project or programme to achieve its desired goal, not only must the product be something that the target population wants and needs, but the delivery system must ensure that they get it. An efficient delivery system may need to be capable of targeting relatively specific subgroups of the population such as women, the poor or the vulnerable. The basic questions that need answering are:

- Do the intended beneficiaries have **access** to this product? (Do they know about it? Is it physically accessible to them? Can they afford it?)
- Do they **use** this product?
- If yes, are they **satisfied** with the product?
- If not, why not?

From these questions, it is then possible to generate three basic indicators:

- ***access* – percentage of the target population having access to the project product.** The term "**access**" has to be clearly defined. It may be "time taken to reach" or "distance" or possibly "ability to pay".

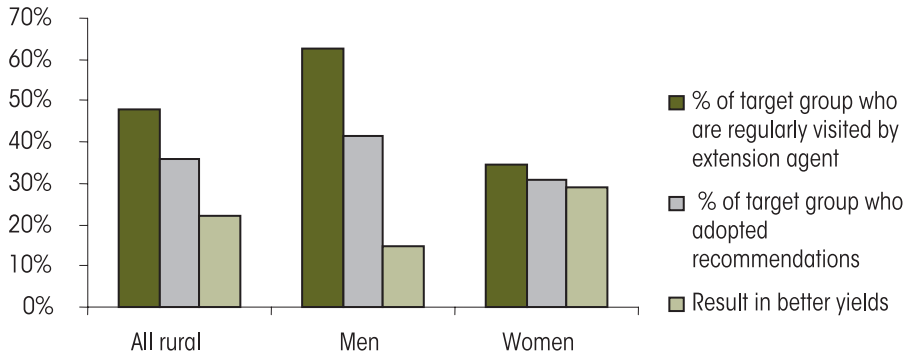
- ***use* – percentage of the target population that uses the project product.** Similarly, the term "**use**" has to be defined. It could for instance be "adoption" as in "percent of smallholders adopting a practice recommended by extension".
- ***satisfaction* – percentage of users satisfied with the product.**

Box 6 shows how these indicators can be applied and adapted to monitor agricultural extension services. Although they are simple indicators, they have a number of qualities that make them attractive as outcome indicators. They are relatively quick to process. This means that the results can be presented very soon after data collection and can consequently be used to sound an alarm in the case of unexpected results.

They can also be collected regularly in order to build up time series, with the first year serving as a baseline. This is important for making before-and-after

Box 6. Adaptation of research and extension service delivery indicators (access, use and satisfaction) to the new Technology Transfer Paradigm

The graph shows how traditional service delivery indicators collected through a household survey of smallholders may be used to monitor the effectiveness of an agricultural extension programme. Access has been defined as "persons having had contact with an extension agent in the last two weeks". Use is defined as "persons who have adopted a set of technological recommendations". Satisfaction is defined as "persons who considered that the recommendations had contributed to higher yields or had otherwise been beneficial". The indicators have additionally been disaggregated by gender.

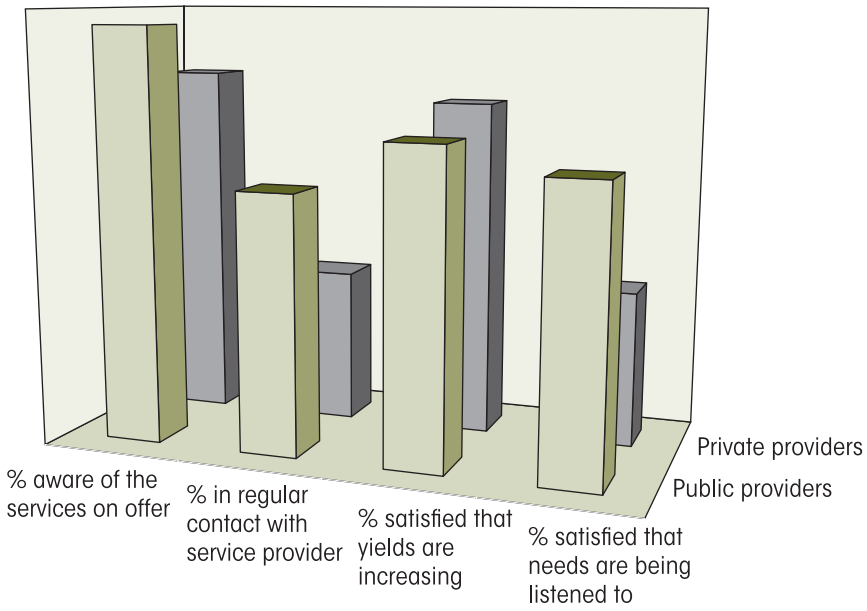


continue

The indicators used in the above example were developed at a time when agricultural extension programmes were based on a view of technology transfer in which farmers are passive recipients at the very end of the innovation process.

This approach is being progressively superseded by the new vision of innovation systems in which farmers, farmers' organizations and communities play a more active part in defining the content of the technology development programme and in which the concept of publicly funded and state-owned extension services is substituted by the approach of pluralistic, public/private, advisory services where farmers choose the service provider and pay for it.

Under such circumstances, the indicators have to be adapted, but the overall service delivery framework can still be maintained. This can be done first by restructuring the questions to the farmers so that a separation is made between the different service providers (public and private) and so that indicators can be separately calculated for each type of provider and second, by recognizing that the "service" is no more just the technological recommendations, but also includes the provision of opportunities for farmers to express their needs. Thus, the satisfaction questions may be expanded to include questions on the extent to which farmers feel that their needs are being listened and catered to.



comparisons. They can also be disaggregated so that comparisons can be made between the answers given by different subgroups of the population (such as by gender, socio-economic group or regional location). They can equally well be aggregated upwards – as long as care has been taken to ensure that consistent definitions are used – so that responses from different countries can be compared at regional and global levels. Nevertheless, a key question needs to be asked: "How easy are they to collect?" There are basically three options: institution-based surveys; community surveys, or household surveys.

Institution-based surveys aim to collect the information directly from or through the institutions that are delivering the product or service, e.g. a fertilizer distribution centre or a rural bank. Reference has already been made in this chapter to QSDSs.

Focus groups or community surveys work at the community level using a community survey with focus group discussions. Using well-trained enumerators to guide the discussions can be very effective in getting people to talk about the project or programme, and at delving below the surface to understand why a service is or is not meeting the needs of a particular user group.

Household surveys will be reviewed in greater depth below, but it can be pointed out here that these surveys are well suited to the collection of service delivery indicators.

A doubt may be raised about the validity of using "satisfaction" as a measure of success. Can one really trust the respondent to give an honest answer? How can one quantify such a subjective notion? There is no reason why a subjective assessment such as satisfaction is not a valid indicator to include among the early measures of outcomes. In fact, who is better suited to evaluate a product than the user him or herself? Monitoring and evaluation are not exact sciences but involve a process of picking up information from various sources and of combining and comparing them to arrive at the most probable assessment. The respondent's opinion is as valid as any other source of information, and although subjective, it can still be quantified. It is generally recommended that independent agencies – not the service providers – should gather the data from the intended beneficiaries so as to reduce possible bias. It can also be useful to collect information both from the service provider and the service user, and to carry out an analysis of the perception gap.

"Satisfaction" is a qualitative concept that can be measured in a quantitative way.

Thus, by employing the service delivery approach, it is possible to set up a system using just a few basic indicators that can serve as a means both to track results and to signal early warnings where results stray significantly from expectations. The service delivery approach works for a large number of projects, including safe water, health care, immunization, electricity, schooling, employment, credit/financial services, roads, public transport, telephone services,

postal services, agricultural inputs and police services. But it does not work in all cases. For instance, it might be difficult to apply it to a component where the main objective was "institutional reform", or to assess the effects of a policy change. Yet even there, questions such as "How has the economic situation of your household changed over the last 12 months?" can provide very useful early indicators of changing circumstances and overall satisfaction with government performance.

In promoting the use of service delivery indicators, there is no suggestion that other measures of project outcomes should be dropped. Production and yield indicators are clearly necessary, but are problematic and long-term. Further, as shown in the next section, it may take a number of years before lessons can be drawn from them. Annex 1 contains a list of suggested indicators relevant to the ARD sector programmes. Some of these may already be available in the country but not collected on a regular basis; others may require collection mechanisms to be established. It is important that systems be put in place to start capturing them early on so that baseline measures can be taken and time series started. These indicators should be taken as a minimum set to which other indicators can be added.

Sector- and national-level outcomes

Up to this point, the discussion has focused largely on M&E of the project level. When it comes to monitoring at the sector level, the principles remain the same. However, the range of products increases and the interaction between programmes takes on increased significance since ultimately, the M&E findings will affect how resources are allocated to each of them. This could lead to the installation of very heavy M&E programmes and to difficulties in coordination.

Sector-level M&E must aim to compare the relative contribution of the different programmes towards the achievement of shared goals.

Fortunately, as one moves up the results chain, one finds that the various projects/programmes are all contributing to the same common goals – the country development goals. The task of monitoring progress towards these goals is no longer a project-specific activity, but a shared one. This calls for a pooling of information and data, and for the standardization of methodology, concepts and definitions. At these higher levels of the results chain, data come partly from the accumulated body of information disseminated through the individual project M&E reports and partly from additional data that will need to be collected. Working at the top end of the results chain is less a question of monitoring indicators than of systematic analysis. It can be a very data-demanding exercise, especially since such higher-level indicators become increasingly costly to collect and complex to analyse. A weak statistical and analytical infrastructure imposes severe limitations on what can be achieved.

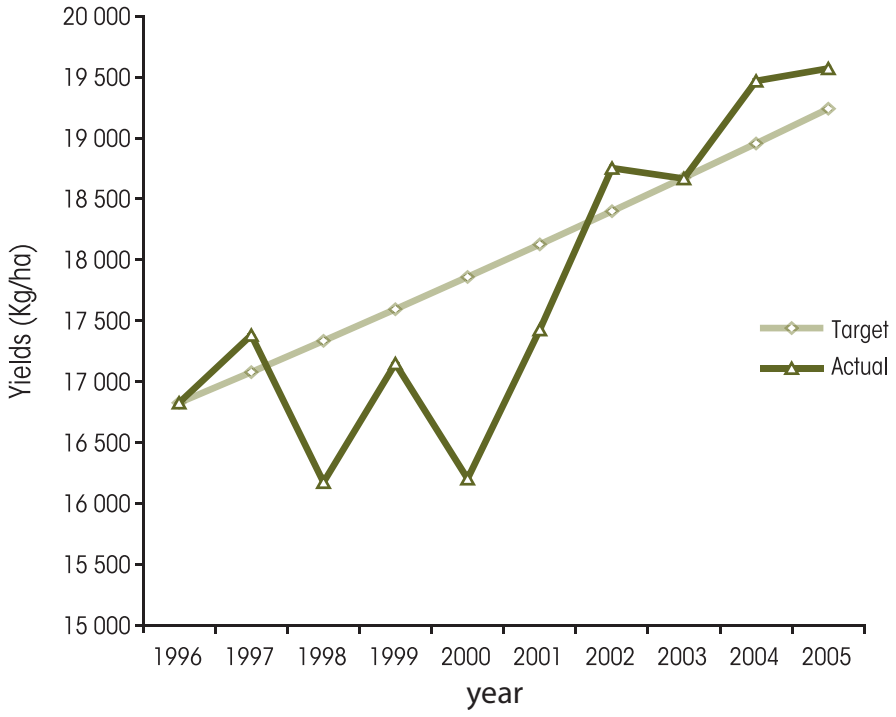
It is not so much that the number of indicators increases, but rather, that complexity increases. Many indicators at this level are quoted as ratios, and separate estimates are needed for both the numerator and the denominator, both of which are potential sources of error and bias. Indicators need to be chosen with care.

Difficulties with the measurement of agricultural output

For monitoring the results of ARD programmes, the most obvious outcome indicators are those that relate to the measurement of changes in production levels (crop, livestock or fish) and yields. While these measures are central to most M&E programmes for the ARD sector, they bring their own particular problems.

Since most agricultural projects share the goal of raising agricultural output, one would think that the simplest indicator would be to measure "yields" – calculated as the ratio of production to area cultivated – and see how they change over time. Unfortunately, it is not that easy, for two reasons. The first reason is essentially a statistical one and centres around the issue of time series analysis. The problem is that agricultural production fluctuates and can vary significantly from one year to the next, primarily but not exclusively due to the

Box 7. Detecting a trend in maize yields



strong effects of rainfall, or the lack of it. This phenomenon is particularly acute in non-irrigated conditions. As a result, it is frequently not possible to detect any change in the trend until a number of years have passed – as many as seven or eight. It is common to see project appraisal documents with projected yield increases similar to those shown in Box 7 (light line). The target is a steady two percent increase in yields per year. This looks reasonable and not too difficult to monitor. But when *actual* yields (dark line) are measured and superimposed over the anticipated trend line, it becomes clear that sharp year-to-year fluctuations in yields make the drawing of any conclusion almost impossible, particularly for the first six years when it would appear that there is no upward trend at all. In this particular case, when the final four years are plotted, the trend line does in fact show an increase of almost exactly two percent a year, as anticipated. But it is statistically impossible to determine this until well past year 6. Random and erratic year-to-year fluctuations of the kind that rainfed crops are prone to experience will severely complicate attempts to carry out time series analysis within too short a period.

But that is not the only difficulty. There is also the problem of measurement errors – errors associated with the measurement of smallholder crop areas and crop production. The classic methodology is to use randomly harvested crop cuts to estimate production and yield. Although this methodology is being successfully applied in many countries, it is known that crop cutting can lead to overestimates of as much as 30 percent in specific situations. Overestimates are due to a number of reasons, including the "boundary effect"; where there is doubt whether a plant is inside or outside the crop frame, it is usually included inside. Overestimates are particularly high in Africa, where traditional plots frequently include multiple crops, irregular planting density and ill-defined, even non-existent, plot boundaries. This makes the application of the crop-cut technique difficult, particularly in less-than-ideal conditions. However, there are other ways of tackling the problem. Methodological experiments to test the viability of alternative ways of measuring production have come up with some interesting and challenging results that suggest that, at least under rainfed conditions, farmers' own estimates may provide substantially cheaper and faster measures of crop production than "objective measures". Indeed, the estimates may even be better.

Farmer estimates may, in some cases, provide cheaper and quicker estimates of production than estimates derived from objective measurements – and with fewer errors.

Methods using GPS for area measurement have the potential of increasing the efficiency of yield estimates in situations where correct estimates of area harvested may not be available. However, in some areas (hilly areas, very small plots, forest areas, etc.) or where plots are irregularly shaped, measurement errors may still be an acceptably high.

These caveats notwithstanding, the measurement of agricultural production will continue to be a central component of any ARD programme, but one should be aware of the potential for error and be on the lookout for alternative ways of assessing results. On the positive side, the introduction of modern farming practices, combined with the arrival of new measurement methods including the use of satellite imagery, are beginning to make the life of the agricultural statistician a little easier. Also, as time series start to lengthen, it becomes easier to identify and discard the obvious outlier years and to reduce the risk of misinterpretation.

The challenge of measuring poverty under less-than-ideal conditions

The ultimate goal of nearly all ARD projects and of the PRS as a whole is to reduce the level of poverty, i.e. to increase rural incomes as a whole and at the same time to reduce income disparities between the rich and the poor. If the measurement of agricultural production was deemed difficult, the measurement of living standards is even more challenging. In order to track the first MDG poverty indicator – "percentage of the population living on less than one dollar a day" – a detailed household survey is required. This may involve multiple visits to households, and the collection and processing of 200 or more items of data from every sample household to compute an estimate of household consumption. Further information has to be provided on all household members, including their age and gender, in order to estimate per capita consumption. More data is then needed on comparative prices before the complex analytical task can begin establishing who is and who is not below the poverty line. In most countries, this is not the kind of indicator that can be realistically measured more frequently than once every five years or so. At the same time, given the close correlation in most countries between household incomes and agricultural production, all the problems associated with the estimation of trend from a time series analysis discussed in the previous section apply equally to the measures of poverty and to the measures of agricultural production. This leads one once again to be on the lookout for alternative measures or methods that could be applied in countries where conditions are less than ideal. Thus, in certain countries, where the goal of regularly monitoring changes in poverty levels may be unrealistic, it may be more productive if instead of focusing on the question "What proportion of the population are below the poverty line?", the analysis focuses on the question, "Are the anti-poverty programmes and services actually reaching the poor and vulnerable as well as the non-poor?" This then becomes an easier question to answer. It focuses attention on the provision of services rather than on the measurement of poverty, but it still requires the classification of households into two classes – the poor and the non-poor. The standard way of doing this would be to establish a national poverty line based on minimum food and non-food requirements, and then establish who is above and who is below this fixed line. This is an absolute measure of poverty, but again, the establishment of such a poverty line can be difficult. An alternative and to some extent simpler solution

is to use a relative concept of poverty. For instance, instead of having a fixed poverty line, one could simply decide to classify, say, the bottom 10 percent as being "the poor". All at once, all the complexities of establishing the poverty line are removed, and the analytical task is simply to compare the services reaching the bottom 10 percent compared with those reaching the rest of the population.

But the problem remains that households must still be ranked using some wealth-correlated variable, such as household income or consumption, which would still require a periodically updated household expenditure and consumption survey. For many countries, this is simply not practicable. However, a number of countries are now experimenting with much lighter household surveys that do not involve the collection of consumption data, but collect specific, easy-to-measure indicators of household well-being. Such indicators may include, *inter alia*, asset ownership, number of literate adults, number of children malnourished, housing quality, mean number of persons per room, and adults unemployed. These are used to create a composite poverty index. Households are then ranked using this composite indicator, and then grouped into deciles. Once this point has been reached, comparisons can be made between deciles. The point is that, even if it is not possible to measure the absolute number of households living in poverty, these short-cut methods allow to identify and isolate those households that are at the bottom end of the distribution, whatever the welfare indicators, and to observe whether they are getting any direct benefit from the various ARD programmes under review.

Evaluation

Finally, one must not forget the "E" in M&E. Monitoring and evaluation are parallel and complementary activities. It is important to be rid of the notion that monitoring is an activity that takes place at the beginning of the project, and evaluation, at the end. Wherever and whenever there is a monitoring activity, there needs to be a regular process of review – of questioning what the data mean and thinking through what the implications are for policy and for the future. Hence, both monitoring and evaluation are continuous activities throughout the life of the project. It is generally thought that evaluation is complex and data-demanding. It need not be so. There are a range of available types and methods of evaluation – programme reviews, interviews with key stakeholders, focus group meetings, performance audits, etc. – that do not require much in the way of additional data, and that can and indeed should be built into the M&E work programme.

Without evaluation, there is no learning; without learning, there is no progress.

What is true, however, is that as one progresses up the results chain, the tasks of evaluation can become increasingly more challenging, and in consequence, require more data. In the early phases of implementation, evaluation may be

no more than the annual review of inputs and outputs to guide the allocation of further resources during the next year. Further up the chain is where the problems lie.

The first task is simply to take the selected outcome indicator and to establish whether it is possible, over a predetermined period of time, to establish a trend. We have already seen how difficult a task this is, particularly where the expected outcome is an increase in agricultural yields. Just establishing a positive trend may require eight or more annual observations. But if this was difficult, then even more so is the task of determining the extent to which the change can be attributed to specific project interventions. The domain of **impact evaluation** and **social policy and impact analysis** will now be discussed. These are analytical tasks that extend way beyond the analysis of simple indicators. Impact evaluation may be undertaken at any level: project, sector or country. Ideally, it requires information on key indicators before (baseline data), during and after the specific intervention or reform. It may involve the setting up of a quasi-experimental design that controls for sample characteristics and permits testing against counterfactual hypotheses so as to compare both the before/after situation and the with/without situation. The complete evaluation should also identify any unexpected or unanticipated outcomes. A full review of impact analysis techniques is beyond the scope of this Sourcebook, but interested readers are referred to Ravallion (2008a and b) for a more complete description of the main methods for counterfactual analysis.

It is important that, where it is assumed that an impact evaluation will be carried out, the expected path that the analysis will take is mapped out as early as possible so that the data requirements can be assessed and addressed accordingly. The process that has just been described for the selection of outcome indicators is in itself a preparation for an impact analysis down the road. It sets out a specific conceptual framework and identifies channels through which the programme/project services are to be transmitted. It is also important that, when selecting the indicators, thought is given in advance to the need to select indicators in such a way that the impact on gender and on the environment can be extracted and evaluated.

What emerges from this is that if careful thought is given at the very start of the project to the selection of indicators to be monitored, and if they are selected so that they catch the most critical stages of the expected transmission mechanisms, then the additional data demands of the evaluation can be minimized.

The burden of evaluation can be minimized in countries with limited resources.

Several lessons emerge for those operating in less than ideal conditions. Not all projects/programmes need full-scale impact evaluations. These should only be conducted where it is thought that there are lessons to be learned. Second, evaluation does not always mean that much additional data is required beyond what has been routinely collected

for monitoring purposes. Third, the additional data needs can be reduced by thinking ahead at the beginning of the programme. Fourth, given the fact that most projects converge towards a single common goal, there are enormous synergies to be gained by looking at certain aspects of the evaluation of impacts at the sector or country level, rather than at the project level. Fifth, if quantitative data are scarce, good use can be made of qualitative studies that can yield valuable and important insights. Finally, where there is clearly a need of serious evaluation, it needs to be planned well in advance, include both qualitative and quantitative studies, and to take into account both expected and unexpected outcomes. It will almost certainly involve combining data from various different sources, and coming to a considered view about the impact of a particular intervention. The benefits of good evaluation are, however, frequently under-appreciated. Evaluative research also has some of the properties of a public good, in that the benefits spill over to other projects. Development is a learning process, in which future practitioners benefit from current research (Ravallion, 2008a and b). The implications of such a research agenda, with respect to the data needs, are considerable.

A CORE SET OF PRIORITY INDICATORS FOR ARD PROGRAMMES

We now complete the work on identifying and prioritizing suitable indicators by bringing together all the indicators that have been discussed so far, and linking them in with the indicators for monitoring national development objectives as specified in the PRS documents.

> ***In order to establish a minimum set of core indicators, a country must comply to international standards.***

We started by noting that there is a difference between monitoring performance and monitoring results. We noted that, for the most part, performance indicators could be monitored using information derived from internal MISs and we looked at some of the tools now available to help improve the monitoring process. Next, we grouped our results indicators into indicators for monitoring early results and indicators for monitoring medium- to long-term results. The early results indicators consisted primarily of service delivery indicators for each of the main ARD products. These service delivery indicators should be supported where possible by quantifiable outcomes, such as yield increases, resulting from target populations adopting or using programme and subprogramme outputs. However, these may need to be tracked several years before any reliable conclusions may be drawn.

There is another set of outcome indicators that is equally important. It covers those that are not directly project-linked – or more correctly, those linked to multiple projects. These include macro- and national-level indicators and indices – the indicators that move as a result of broad policy changes or of the combined effects of several programmes or interventions. They include price indices, food production, agricultural exports, fertilizer use and imports. They also include

some of the more common multi-sectoral indicators that may be used to compare the rural and urban areas, and to measure the results of the combined package of policies and interventions specified in national development strategies. Examples of these include: the proportion of population living in poverty, GDP per capita; urban/rural comparisons of multi-sector indicators such as prevalence of underweight children under five years of age; ratio of girls to boys in primary and secondary education; and the proportion of the population with sustainable access to improved water sources.

The process of selecting a comprehensive set of indicators that meets everyone's requirements is not easy, since different users at different levels have varying information needs. Ideally, the process of selection should be participatory and take into account the needs of all stakeholders, and the principle should be retained that countries select their own indicators according to the content and goals of their PRSPs. The process can be facilitated, however, by drawing on the experience of what other countries have done. Annex 1 provides a menu of indicators that countries can use to help them prioritize and select the most useful indicators for their particular needs. The list is not exhaustive nor is it expected that all countries should adopt and use all of them. Some may not be relevant and others may lack the country capacity to collect them, but the list offers a choice and includes examples of good practices taken from different countries around the world. The indicators include measures of early results as well as medium- to long-term results. They are provided for all the main ARD subsectors and related themes, and countries can choose which ones to use.

The priority indicators need to be underpinned by a database of core ARD statistics.

For monitoring ARD goals at the international level, however, there has to be standardization. A subset of 19 essential indicators have been identified from among the full list and labelled as priority indicators. Some of these indicators already appear in the FAO statistics database (FAOSTAT), but for many countries, the series are either non-existent or incomplete, with significant gaps or with the values that have been filled by imputation. The international series are in need of urgent upgrading, but the quality of the series can only be improved if all countries commit to maintaining the same indicators at national level, and agree to adhere to common standards. These priority indicators represent a minimum core set that all countries need to maintain and update on a regular basis. Without this minimal commitment at the country level, it is not possible to improve the quality of M&E at the international level. But this should not be too onerous a burden, since the same indicators serve not only to monitor at the international level, but also at a national level. The priority indicators on their own are not enough to meet all M&E data needs, but they should be seen as an essential subset, and as far as possible, they should be included in all national M&E programmes. The priority indicators are shown in Box 8 and the expanded list of indicators are found in Annex 1.

Box 8. List of priority indicators

A Sector-Wide Indicators for Agriculture and Rural Development	
Early outcome	
P1	Public spending on agriculture as a percentage of GDP from the agriculture sector.
P2	Public spending on agricultural input subsidies as a percentage of total public spending on agriculture.
P3	Prevalence (percentage) of underweight children under five years of age in rural areas.
Medium-term outcome	
P4	Food Production Index.
P5	Annual growth (percentage) in agricultural value added.
Long-term outcome	
P6	Rural poor as a proportion of the total poor population.
B Specific Indicators for Subsectors of Agriculture and Rural Development	
1. Crops (inputs and services related to annual and perennial crop production)	
Medium-term outcome	
P7	Change (percentage) in yields of major crops of the country.
2. Livestock	
Medium-term outcome	
P8	Annual growth (percentage) in value added in the livestock sector.
3. Fisheries and aquaculture	
Long-term outcome	
P9	Capture fish production as a percentage of fish stock (or a rating of the state of major capture fish stocks relevant to exports and local food).
4. Forestry (developing, caring for or cultivating forests; management of timber production)	
Long-term outcome	
P10	Proportion (percentage) of land area covered by forest.
5. Rural Micro and SME Finance	
Early outcome	
P11	Percentage of the rural population using financial services of formal banking institutions.
6. Agricultural Research and Extension	
Early outcome	
P12	Public investment in agricultural research as a percentage of GDP from the agriculture sector.
7. Irrigation and Drainage (services related to water use in agriculture)	
Early outcome	
P13	Irrigated land as percentage of crop land.
8. Agri-business (agricultural marketing, trade and agro-industry)	
Medium-term outcome	
P14	Change (percentage) in sales/ turnovers of agro-enterprises.
C Indicators for Thematic Areas related to Agriculture and Rural Development	
1. Community-based Rural Development	
Early outcome	
P15	Percentage of farmers who are members of community/producer organizations.
2. Natural Resource Management	
Medium-term outcome	
P16	Withdrawal of water for agricultural as a percentage of total freshwater withdrawal.
P17	Proportion (percentage) of land area formally established as protected area.
P18	Change (percentage) in soil loss from watersheds.
3. Land Policy and Administration	
Early outcome	
P19	Percentage of land area for which there is a legally recognized form of land tenure.

The exercise of validating identified indicators at the country level was aimed at testing the "relevance" of the indicators to the current development activities and the feasibility of their compilation in less-than-ideal conditions.

In recommending the 19 priority indicators, greater attention has been given to the criteria of "comparability" across countries and "availability" of data for their compilation, in addition to "relevance".

Box 9. Cambodia's two-tiered system

The development of the national M&E system in Cambodia is anchored on the country's National Strategic Development Plan (NSDP). The plan is a single, overarching document containing the priority goals and strategies of the Royal Government of Cambodia to accelerate the reduction of poverty and to achieve other Cambodian Millennium Development Goals (CMDGs) and socio-economic development goals for the benefit of all Cambodians.

The M&E system adopts the "two-tiered structure" as its operational framework. It consists of a set of performance indicators, derived from the framework and the priorities of the NSDP, together with effective mechanisms for tracking progress. It aims to ensure regular and periodic M&E of the provision of inputs, achievement of outputs and outcomes of various strategies and actions under the NSDP.

At the national level (first tier), a limited and manageable number of 43 core indicators have been selected. These are aligned with macro-development goals and targets to achieve CMDGs. These are also used to monitor key dimensions of NSDP progress, and provide the basic framework on which annual progress reports are prepared.

The second tier is used at the line ministry/agency level. Each line ministry/agency is required to develop its own set of performance indicators using CMDG indicators (referring to the 43 NSDP-based core indicators) under its jurisdiction, and other indicators relevant for sector-level monitoring purposes. The aim is to create a more in-depth and disaggregated picture of the ministry/agency-level support to detailed policy/programme monitoring and analysis, and reorientation. Guided by the NSDP, the development and selection of indicators at the line ministries/agencies should:

- facilitate informed decision-making and help re-set priorities and policies;
- enhance transparency and accountability through improved information sharing;
- promote a better understanding of the linkages between NSDP implementation and resulting outcomes.

Box 9 describes how a process very similar to the one described here was used in Cambodia in the selection of indicators for monitoring their PRS.

It is not enough, however, to simply develop a list of desirable indicators without at the same time identifying the data that will be needed to calculate them. Thus, linked to the concept of priority indicators is the idea of maintaining a set of core statistics data series needed to underpin the indicators. Once these statistics are added together, the modest list of data requirements starts to grow very quickly, with significant implications for the NSS. This "shopping list" of data needs provides the basis for a dialogue with the suppliers. For most of the outcome indicators, the supplier will be the NSO. It may also include other agencies that make up part of the NSS. The objective of the dialogue is to negotiate arrangements for a programme of survey activities that will ensure the delivery of the appropriate data according to the timeline specified. This is the subject of the next chapter.

Chapter 3
The data framework

When M&E specifications are being established, it is often not taken into consideration how expensive and resource-consuming the process of data collection and dissemination can be. It is at this early planning stage that overambitious expectations can lead to the creation of an M&E programme, which, because of its complexity, has little hope of success. This chapter looks specifically at the issue of data supply and reviews various tools and approaches that have been used with some success in different countries. The chapter concludes with a discussion on the capacity of a National Statistical System to support M&E data needs.

It is clear from the previous chapters that even the lightest of monitoring systems can make extensive demands on the data supply system. In order to meet the needs of monitoring at each of the four levels (inputs, outputs, outcomes and impact), the M&E system needs to draw on information coming from a variety of different sources. It is not just that each level requires different indicators, but also that the requirements in terms of periodicity, coverage and accuracy vary according to the level of indicator. Input indicators are required to inform short-term decision-making. They therefore need to be produced frequently and regularly – possibly once every 1-6 months. The same applies to output indicators, but here the reporting period can likely be longer, say, one year. As one moves further up the results chain and starts to collect more information about clients rather than the servicing institution, the task of data collection becomes more complicated, the tools less reliable, and the results more questionable. To counteract this, it is advisable to use information from different sources and to use different methods to arrive at a reasonable estimate of the outcome under review. On the other hand, the time frame can be relaxed – a little. Time must be allowed for clients to become aware of and start using public services. One may see little evidence of outcomes for the first few years. Therefore,

M&E systems need to draw on a wide range of information sources. Baseline information is important for evaluating with and without project effects.

it may be acceptable to build a programme around the reporting schedule of, for instance, 1-2 years. But it is important that the process is initiated at the very beginning of the project with a view to using the first report for establishing the baseline situation. The evaluation of the eventual impact comes much further down the line – often years after the project has been completed. Although the time frame may be more relaxed, the analytical challenge is not, and from the data collection perspective, experience has shown that it is vital that the outline on how the project is to be evaluated is agreed from the very beginning, since it may involve setting up an experimental design to try to isolate the "with/without" project effect.

So, what is available to support the establishment of simple but effective M&E operations? What tools are available? The following list is not comprehensive, but each supports a different part of the M&E jigsaw puzzle. They include different types of household surveys, rapid appraisal and participatory methods. All are used to provide the necessary data for the calculation of the "upper end" indicators, namely outcomes and impact indicators. They include both quantitative and qualitative assessment tools.

The tools

Household survey elements

The most popular and obvious instrument for monitoring the outcomes of ARD programmes and the contribution made to poverty reduction through ARD is a household survey. There are other options, of course. If we review the list of results indicators shown in the previous sections, we see there is a possibility of collecting basic data using administrative records, community surveys or even individual focus group interviews. All have their strengths and limitations. But the great strength of the household survey is that it provides information both on the beneficiaries and on the non-beneficiaries. It also has the advantage that the indicators derived from the survey can be both aggregated and disaggregated to different levels. It can thus serve as a tool for monitoring at the global level as well as at the national and subnational levels.

The great strength of the household survey is that it provides information both on the beneficiaries AND on the non-beneficiaries.

The distinguishing features of a household survey are that it uses a fixed format questionnaire, which is administered to a probability-based sample of respondents who represent a particular population (usually the intended beneficiaries of the programme – the clients).

Sample

Statistical surveys use random sampling to ensure that the information collected is unbiased and that the size of the error that may result from using a sample rather

than a complete enumeration is known. Clustering facilitates survey fieldwork and logistics but reduces the sample efficiency. This can be partly compensated for by stratifying the clusters into homogeneous groups before the selection is made.[1]

The question is often asked, "How big should the sample be?" In the textbook approach to sample size determination, size is determined by the variability of the characteristic of interest, the way in which the sample has been designed and the degree of precision that the user needs.[2] For practical planning purposes, however, a very rough but frequently used rule of thumb is to think in terms of a sample size of 500 to 600 households for each analytical domain, i.e. the subgroup of the population for which indicators are required. Sampling errors diminish as sample size is increased. It is evident, however, that since the requests are made for increasingly lower levels of disaggregation, sample sizes quickly increase to unmanageable proportions. This is one of the trade-offs that has to be considered when designing a survey.

Planning a survey is all about trade-offs.

Questionnaires

The second key characteristic of a household survey is that it uses a structured questionnaire in which respondents' answers are recorded. A questionnaire with a fixed format allows data entry into a structured database, with a minimum amount of manipulation, so that it is ready for validation and analysis. Good survey practice dictates that questionnaires should be printed in the same language in which the interview is to be conducted, but in many developing countries, there may be 20 to 60 or more local languages, making it impractical to translate in all languages. This introduces the concept of "non-sampling errors", which are all the errors that can occur during the course of the survey that are not related to the sample or sample design. Unlike sampling errors, whose size can be mathematically calculated, the magnitude of non-sampling errors is generally not known, but it may be safely assumed that they are significantly greater than those of the sampling errors. In contrast to sampling errors, which decrease in size as the sample is increased, non-sampling errors have a tendency to increase with sample size. This is another trade-off that has to be considered in survey planning. In principle, the wisest course of action may be to consider and plan for minimizing non-sampling errors when preparing the overall survey design, and build checks and balances into the survey and data handling processes.

Survey design

A third feature of household survey is the survey design. This includes all the survey logistics, the numbers of visits to be made to the households, the reference

1 Typical stratification criteria include urban/rural clusters and/or stratification by agro-ecological zone.
2 Note that sample size is not a function of population size; the common belief that the size of the sample should be a certain percentage of the population is therefore misconceived.

periods that will be used in the questionnaire and the choice of which household member or members are to be used as respondents, etc. These are often the factors that distinguish most clearly one type of household survey from another. Even minor changes in design from one round to the next can have significant effects on the results. This introduces the degree of conservatism in the NSOs, which, being unwilling to disrupt time series, may resist change. However, for the purposes of making global comparisons between countries, it presents some limitations. The problem is not considerable with simple indicators such as anthropometric measurements where the methodology is relatively well established and common across all countries; it is a problem, however, with complex computed variables such as household consumption, another primary poverty measure used for tracking the first Millennium Development Goal. A third set of trade-offs to be considered, therefore, are the relative advantages and disadvantages of using a nationally developed methodology compared to a standardized international survey design.

Data processing, storage and dissemination

Nowadays, good survey practice highlights the fact that data processing involves not just the tasks of data entry, processing and table production, but goes much further to include data storage and archiving, and electronic data dissemination. It also includes the storage, archiving and dissemination of *metadata* together with the actual data. The complete survey package can fit neatly onto one CD, which can be readily disseminated and made available to users. One issue that continues to concern many countries is the question of a **data access policy**. In many countries, access to survey data remains highly restricted. Confidentiality is often cited as the rationale, but the real reasons are often political or organizational. Users may be granted access to the data in aggregate form, but for many practical purposes, this is not enough; they need it at the unit (household) level. It is therefore important that, right from the start, clarity be achieved as to what the data access policy will be. Through the International Household Survey Network sponsored by the World Bank, United Nations agencies and regional banks, tools for documenting and disseminating microdata according to international standards and practices have been developed and country capacity is being strengthened with the support of World Bank/PARIS21 Accelerated Data Program (see www.internationalsurveynetwork.org/home). Also, FAO has developed the CountrySTAT system as an integrated platform for better harmonization, access and dissemination of country-level food and agriculture statistics (www.fao.org/statistics/countrystat).

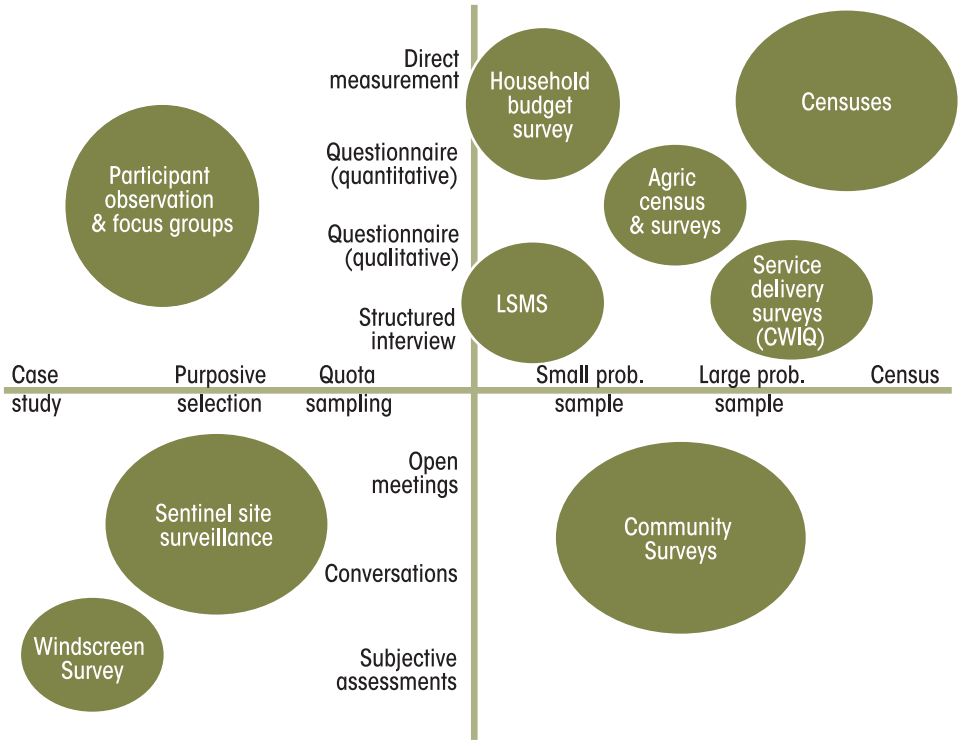
Questions about data access need to be addressed at the very start.

Different household survey models

Household surveys can differ widely: different models serve different purposes. Box 10 highlights some of the different ways of collecting information from households, including both qualitative and quantitative approaches. It plots the most commonly used surveys on two axes. The vertical axis – the qualitative/quantitative axis – represents a range of different methodological approaches from subjective assessments through to direct measurement. The horizontal axis shows different levels of representativeness, from the simple case study (not representative) right through to the population census, which is fully representative. Different types of surveys have been superimposed onto these two axes, where they can be seen to scatter from the lower left-hand corner (non-representative/subjective) up through to the upper right-hand corner (fully representative/objective). This helps to decide on the right instrument for the task in hand.

Most of the statistical surveys are to be found in the top right-hand quadrant, whereas the more qualitative studies tend to be clustered in the lower left-hand quadrant.

Box 10. Tools for measuring results: surveys vs. non-formal appraisal methods



Population census

The population census appears in the top right-hand corner. It uses a short questionnaire, which should be administered once every ten years and should cover the entire population. Its value lies not just in the fact that it provides a complete account of every person in the country, but that it also serves as a basis for nearly all subsequent sample survey activities.

Relevance to monitoring ARD programmes: The census is pivotal to any survey programme. The census results plus the cartography work conducted beforehand provide essential information for preparing sample frames for any subsequent sample surveys. When combined with household survey data, census information can be used for the creation of poverty maps and atlases of social indicators.

Duration: Even though fieldwork may only last a few weeks, there is an enormous amount of preparatory work – two or more years – leading up to census day. Preliminary results, in terms of simple cross tabulations and counts, can usually be made available within a few weeks of the end of fieldwork. Full results are often not forthcoming for a year or more, however, and require clearance at the highest political level.

The population census is pivotal to any survey programme. When combined with household survey data, census information can be used for the creation of poverty maps.

Questionnaire size: The size should be three to four pages. There is usually little opportunity to add substantive questions, but it may be possible to include a few socio-economic classification variables such as "Does the household operate a holding?".

Cost: Censuses costs vary enormously, but a commonly used rule of thumb is to work on the basis of one dollar per person. Thus, for a population of 10 million people, the cost of census would be approximately US$ 10 million.

Agricultural census and agricultural surveys

The agricultural census: Closely associated with the population census is the agricultural census. FAO recommends that an agricultural census be conducted at least once every ten years, just as the population census. The new World Programme for the Census of Agriculture (WCA) 2010 advocates a system of integrated agricultural census and surveys, and introduces a modular approach. For the core module covering 16 data items, a complete enumeration is recommended, while for supplementary modules, sampling can be used. The new programme shows how integration of an agricultural census with a population census and other agricultural surveys could prove cost-effective and enhance the scope of data-analysis. The traditional role of the agricultural census as a provider of structural

data at the small geographical level has been amplified in the WCA 2010 to view it as a vehicle for monitoring the MDGs and other ARD policies. Recognizing the increasing demand for community-level data in the development planning and monitoring process, the new programme advocates its collection as part of the agricultural census as well. The 33 suitable data items at the community level presented in the programme include socio-economic aspects of the community as well as access and use of community agriculture-related infrastructure, which may provide useful information for planning and impact measurement. The programme provides an option to the census planners to widen the scope of the agricultural census to cover all the rural households, thus opening up a vehicle for collection of data for monitoring rural development. Data on a number of proxy variables for ARD monitoring could easily be derived from the agricultural census data.

Agricultural surveys are extremely important since they are frequently the only means of monitoring changes in crop production levels and yields. They can also include information on service delivery.

Agricultural surveys: Agricultural surveys may feature as part of the NSO's household survey programme or may be conducted separately by the Ministry of Agriculture. Both arrangements are common. Many countries regularly undertake annual agricultural surveys separate from household surveys for crop forecast and estimation of post-harvest production. In other countries, where they are part of the household survey programme and conducted by the NSO, the trend has been to merge the collection of agricultural statistics with the collection of other household-level statistics using integrated household surveys. Such integration does reduce the cost of data collection and provide some advantages to the analyst wanting to look at the household and holding holistically. There are also disadvantages, however, particularly because the sequence of enumerator visits to the household for integrated surveys makes no allowance for the fact that the collection of data on agriculture should be linked to the agricultural season. For a number of reasons, the quality of agricultural statistics has declined in many countries over the past decade or so, and one of the reasons may be the merging of agricultural surveys with multi-topic household surveys. There is a need for increased priority and more methodological research in this area. This includes the need for more research on such issues as the estimation of agricultural areas and production, not just for different crop types, but for other outputs such as livestock and livestock products, and the establishment of best practices and standards.

Sample size: Sample sizes vary enormously. Agricultural census/surveys are particularly vulnerable to the dilemma that, on the one hand, there is enormous demand for increasingly disaggregated agricultural production data – which

implies large samples – while, on the other hand, current practices for measuring areas and estimating production are slow, cumbersome and prone to significantly larger errors – which implies using smaller samples in order to control non-sampling errors. The increasing use of new tools such as the global positioning system (GPS) for crop area measurement is considerably reducing the work load and cost of this task.

Relevance to monitoring ARD programmes: Agricultural censuses and surveys are extremely relevant since they are frequently the only means of monitoring changes in crop production levels and yields, which are among the key output indicators defined in earlier sections. It should also be noted that both the agricultural census and agricultural surveys may be used as vehicles for collecting data on service delivery as done in some countries (see, for example, the Tanzanian Agricultural Census). The decline in the quality of agricultural statistics must be taken very seriously, being an area in which resources for capacity building are most needed.

Living Standards Measurement Study (LSMS) Integrated Surveys

In the same quadrant of Box 10 but using smaller samples, one finds Integrated Surveys. They are multi-topic surveys that include questions on nearly all aspects of household socio-economic conditions. They may take several forms, one of the best known of which is the Living Standards Measurement Study (LSMS), developed in the 1980s by the World Bank as a data-gathering instrument to conduct research on living standards and poverty. The LSMS uses a large questionnaire filled out in the course of two visits to the household, spaced two weeks apart. During the first visit, the enumerator collects information about all the individual members of the household. This includes information on their health, education, employment and earnings, and on household assets. During the second visit, questions focus on household consumption and expenditure, farm and non-farm enterprises, and earnings. Anthropometric measurements are also taken for all children under five years old.

Integrated surveys are good as baseline surveys: they can measure poverty levels, identify potential problems in need of attention and generally understand the way in which households operate.

Sample size: Because of the size of the questionnaire and the need to control non-sampling errors, sample sizes are generally kept low. Initially, LSMS surveys used samples of 2 000 to 3 000 households, but with the increasing demand for poverty monitoring, sample sizes grew to 8 000 or more households. Even with these larger sample sizes, survey results should still only be presented at relatively high levels of aggregation, such as for urban and for rural areas.

Duration: Fieldwork normally lasts for one year and is carried out by mobile teams of enumerators. Households visits are spread evenly throughout the 12 months. This is good for removing biases in the consumption data, but is, in general, not the most efficient way of collecting agricultural data (see above).

Cost: Integrated Surveys are expensive and may cost around US$2 million.

Relevance to monitoring ARD programmes: LSMS/Integrated Surveys are particularly good as baseline surveys that can be used to measure poverty levels, identify potential problems in need of attention, and generally understand the way in which households establish mechanisms to cope with difficult living conditions. The big disadvantage is that they are difficult to undertake, and if they are to provide baseline data, they truly need to be initiated a year or more in advance of the actual programme. In addition, many countries have neither the analytical nor the survey capacity to successfully carry out such large-scale complex surveys.

Household budget surveys

Household budget surveys are traditionally undertaken to update the basket of goods and services, and recalculate the weights for the Consumer Price Index (CPI). They are more focused than integrated surveys, and the main topics relate to household income expenditure and consumption. But it is rare nowadays not to find a household budget survey that also includes a minimum set of questions on the socio-economic characteristics of household. The line between household budget surveys and integrated surveys can therefore be fuzzy. Because the main area of interest is household consumption, the number and frequency of visits to the household is usually higher than with Integrated Surveys, and the assumption is that the accuracy of the consumption measure will be greater with household budget surveys than with integrated surveys.

Relevance to monitoring ARD Programmes: Household budget surveys are used in many countries as the primary vehicle for establishing and monitoring poverty levels. If they are linked to a light, multi-topic indicators survey such as the Core Welfare Indicators Questionnaire (CWIQ), they can serve a purpose similar to that of an integrated survey.

Service delivery surveys

Service delivery surveys appear in the same quadrant but lower down. They are relatively recent additions to an NSO's repertoire of surveys, but have been used in market research for a long time. A good example of a service delivery survey is the Core Welfare Indicators Questionnaire (CWIQ) (Box 11).

Service delivery surveys are very well-suited to monitoring early results – They are easy to implement and can be repeated annually.

Box 11. Core Welfare Indicators Questionnaire (CWIQ): a survey instrument for collecting service delivery indicators

The CWIQ is a survey tool for monitoring simple indicators and measuring the performance of a range of development programmes. The CWIQ shows who is and who is not benefitting from actions designed to improve social and economic conditions. The CWIQ collects indicators of household well-being and indicators of access, usage and satisfaction with respect to the community and other basic services.

The CWIQ is designed to be administered to large samples of households so that results can be disaggregated to relatively low levels, and to be repeated annually so that time-series can be quickly built up. It is intended to complement rather than replace other surveys. It can serve as an annual "core" questionnaire for a National Statistical Office (NSO) to use in a "core and rotating module" survey programme. As such, the CWIQ can become one of the components of a country's overall poverty monitoring package. NSOs should be able to implement the core questionnaire easily each year and add special modules if desired, such as a labour force module or a crop forecasting module.

The CWIQ draws extensively from market research practices and past household survey experiences, as well as recent developments in data entry and processing. As a result, it is a relatively high-tech instrument, but one which requires little in terms of high-tech equipment or training.

The CWIQ focuses on simple indicators of usage, access, and satisfaction. For example, in the education sector, access indicators include distance to primary schooling; usage indicators include primary school enrollment rates; and satisfaction indicators are based on opinion questions to indicate household rating of the quality of services of the current year compared to the previous year.

It also collects a few indicators of household well-being: percentage of households reporting diminishing or increasing assets (land and livestock); percentage of literate adults; percent of children malnourished; housing (quality and mean number of persons per room); percent of adults unemployed in the past four weeks, among others. These are used to create a poverty index, which is later used to rank households and group them into "poverty quintiles". It is thus possible to compare poor with non-poor households.

The CWIQ is an off-the-shelf survey with a number of features designed to improve both the quality and speed of delivery of results.

continue

Simple reporting of results: The CWIQ facilitates the production of a set of standard outputs disaggregated by urban and rural poverty quintiles almost automatically. This allows for quick comparisons between poor and non-poor households in both the rural and urban areas. Data can be easily exported into any of the standard statistical packages for a more rigorous customized analysis.

Large samples: To present and compare social indicators across different population subgroups, the CWIQ should use as large a sample as the local statistical resources are capable of handling. For national surveys, sample sizes of between 5 000 to 15 000 households would be recommended in most African countries. Countries that already have master samples would be in a better position to move ahead more quickly with the survey.

Easy data collection: The CWIQ is based on a single visit to each household only. Because of the simple format and short questionnaire, the CWIQ can be conducted by a non-statistical organization.

Short questionnaire: The questionnaire is four pages long (eight sides).

Quick data entry and validation: The questionnaire uses multiple choice questions and optical mark recognition (OMR) for data entry. Scanners make it possible to enter and clean the data of more than 300 households a day. Basic validation checks are carried out at the same time as data are entered, after which predefined tables and graphs are automatically generated.

Relevance to monitoring ARD programmes: Service delivery surveys are very well-suited to monitoring early results: they are easy to implement and can be repeated annually without disturbing any other survey work that the NSO may be undertaking. Once the questionnaire has been adapted to meet the special needs of a particular country, it is relatively easy to adapt the data processing system so that the processing, storage and dissemination of results can be handled by the NSO with relatively little external assistance.

Other forms of enquiry

Participant observation and focus group discussions

The lower left-hand quadrant contains a wide range of qualitative surveys and studies. These are characterized by the fact that they use small, often purposive (rather than random) samples and do not use fixed questionnaires, but instead rely on relatively unstructured conversations and interviews for the data.

The basic idea is to provide an environment in which respondents share their own views with the interviewer without being fettered by the limitations of a formal questionnaire. These kinds of qualitative studies are sometimes considered to be in competition with quantitative approaches, but they are actually complementary.

Relevance to monitoring ARD programmes: A good M&E system uses a wide range and variety of learning tools to better understand the needs and behaviour of the population that the programme is designed to serve. Quantitative and qualitative approaches can be applied iteratively. For instance, the results of a service delivery survey for an agricultural extension programme may indicate a problem with respect to low adoption rates of recommended practices by a particular class of farmer. It flashes an early warning signal that adoption rates are below expectations, but it is not particularly good at saying *why* they are low. This is often where a few select focus group interviews can come up with a possible explanation quickly and cost-effectively. Such insights often need to be explored further. For example, during the course of the focus group interviews, the suggestion may be put forward that the adoption rates are low because extension agents do not visit lower income households. While this may be true for the participants in the focus group interview, how universal is the problem? The group discussions cannot answer this question, but the service delivery survey could do so with the addition of just one or two simple questions.

Qualitative studies can provide insight into the motives and coping strategies of different target groups.

The Windscreen Survey and other rapid appraisal methods

The Windscreen Survey appears at the bottom left-hand corner of the figure in Box 10. This is really not a methodology at all: it consists of the investigator driving around the project or programme area and observing what is going on through the windscreen. It is more akin to journalism than to serious investigation, but is cheap and quick, and does have a role to play. In Ghana, for instance, forecasts for the forthcoming cocoa crop were made on the basis of expert assessment; the expert in question viewed the crop as he surveyed a wide area by vehicle. Windscreen Surveys can be made more credible by establishing a route that is repeatedly followed over time, supplemented by some simple counts of fields and quality assessments of crop conditions such as "very good", "good", "average", "poor" or "very poor". Rapid assessment techniques should not be dismissed as a source of information as long as they are used in tandem with other methods. They are particularly

The Windscreen Survey is cheap and quick, and can provide useful information.

effective as early-warning devices and can make a significant contribution towards the monitoring of ARD projects and programmes, and can provide important insights if conducted by a knowledgeable expert.

Community surveys

Like household surveys, a community survey can be conducted both with probability and non-probability samples, and can, in principle, be found on any of the four quadrants in the chart in Box 10. For the purposes of M&E, however, it is more probable that they will have the characteristics of surveys located in the lower right-hand corner – relatively representative but subjective. A community meeting is called (usually by the community heads) and certain leading questions are addressed by the enumerator to the community at large. Occasionally, the community survey is directly linked to, and carried out at the same time as, a household survey. The LSMS, for instance, includes a community questionnaire, administered in each sampled cluster at the same time as the households are being interviewed. Its purpose is to collect information about the community and the environment in which the sample households reside. Such information is collected at the community rather than the household level, because the answers will be the same for all households in the community. The focus of analysis tends to be directed towards an examination of the relationship between the household and the community – a *micro-meso analysis*.

Community surveys are particularly good for monitoring community- driven development projects. They can actually become part of the project and owned by the community.

The other way of conducting a community survey is to use it as an alternative rather than a complement to the household survey. In such cases, the unit of analysis is the community itself. In addition, the focus of the analysis tends to be on the relationship between the community and the country as a whole – a *meso-macro analysis*.

The new World Programme for Census of Agriculture (WCA 2010) also includes recommendations for collecting community level data during the agriculture census where appropriate.

Community surveys may be used to collect information on the communities' physical and social capital. They may also be used to collect service delivery information at the community rather than household level. In fact, in countries where the statistical infrastructure is particularly weak – such as in a post-conflict situation – a community survey may be the best way of rapidly assessing what public services are most needed and where.

Relevance to monitoring ARD programmes: Community service delivery surveys can, in the right circumstances, substitute for household service delivery surveys.

Box 12. Nigeria's community service delivery survey

	ACCESS	USE		SATISFACTION		
	a. How far away is this service?	b. How much does your group use this service?	c. Who owns or runs this service?	d. If used, how satisfied are users with the quality of the service provided?	e. Do you have any concerns with the service?	f. How has the quality of service changed in the last 5 years?
	Don't know (Skip to next row) Within community Less than 30 minutes away 30-60 minutes away Over one hour away	Not used at all (answer C then skip) Not as often as needed (go to c,d,e,f) As often as needed / regularly (go to c,d,e,f)	Ownership of Service	Dissatisfied Neither Satisfied or Dissatisfied Satisfied Don't Know	Too far away Too expensive Poor Service Limited staff/ equipment Inadequate building Other_____	Don't know Worse now Same as before Better now
1. Day Care Service	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
2. Pre-Primary service	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
3. Primary service	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
4. Secondary service	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
5. Hospital	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
6. Health Centre (Outpatient)	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
7. Patent/Dispensary Services	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
8. Pre/post natal care	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
9. Immunization/Vaccination	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
10. Mobile telephone	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
11. Post office	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
12. Public transport	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
13. Agric. extension services	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
14. Farm inputs	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
15. Commercial bank	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
16. Police Station	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐
17. Customary Court	☐☐☐☐☐	☐☐☐	[__]	☐☐☐☐	☐☐☐☐☐☐	☐☐☐☐

They are also particularly effective for monitoring community-driven development projects, because the survey can actually become part of the project, and the responsibility for its monitoring can be progressively passed on to the community itself. One of the big advantages of a community survey is that a relatively large number of communities can be covered in a relatively short time. Box 12 shows an example taken from Nigeria of part of a community questionnaire containing service delivery information. It illustrates how a standard set of questions can be applied to a range of different services.

A potential weakness of the community questionnaire approach is that the definition of a community is often difficult to pin down, particularly in urban areas, and it may not be feasible to use probability sampling to select the communities to be interviewed. Therefore, they may not be statistically representative, a problem that most qualitative studies face.

Institution-based surveys

Reference has already been made in Chapter 2 to QSDSs as a means of looking at service delivery issues, but from the suppliers' perspective. One can also use the institution that is supplying the service as a contact point for collecting views on the service user. The principle of collecting information from clients while they are actually making use of the service is common private sector practice, particularly in establishments such as restaurants and hotels. Take, for example, the short evaluation questionnaires on which the guest is asked to rate the quality of service. The problem with such questionnaires is that they are voluntary and therefore only likely to be filled in by people with particularly strong views; the results are unlikely to be representative of the target population. Also, this method provides no information about non-users, which means that there will always be problems in calculating percentages because the denominator is not known. Although not very often used in a development context, variants of institution-based service delivery questions may be observed in some sector information systems, such as in health and education. For instance, information gathered in an annual school census conducted by a Ministry of Education can be used to calculate such indicators as primary school enrolment, which is essentially a usage indicator of the education service. Another more promising way of introducing institution-based service delivery monitoring would be to use institutional administrative records to identify service users who could then be asked to complete a questionnaire. One example might be a livestock-dipping centre. Administrative records will automatically record the number of livestock dipped, vaccinations provided, etc., but these could be supplemented at very little extra cost with service delivery information collected from the livestock owners, using a simple exit poll.

Satellite imagery and aerial photography

Satellite imagery is becoming increasingly accessible, and its resolution has improved to the point that individual fields are relatively easy to identify. The use of imagery is unlikely to replace field surveys (ground truthing is still required), but it can be added to the arsenal of tools for monitoring and evaluating agricultural development. Satellite imagery is also useful in developing sampling frames and as a basis for surveys. The methodology of sampling is now well developed and is in the arsenal of tools advocated by FAO. With some simple procedures, one can mount a household survey using point sampling without the expense and time involved in using a register. Some of the more interesting recent breakthroughs in poverty monitoring include the combined use of imagery, census data and household survey data, which together can be used to create dynamic poverty maps showing changes to key variables over relatively short time periods. Satellite imagery can also be used in developing area sampling frames as basis for area-based surveys, including point sampling. With some simple procedures, a household survey using point sampling could be designed that could minimize expense and time as compared to list frames.

Box 13. Comparison of key features of different surveys

	1	2	3	4	5	BEST USED FOR		
	SAMPLE SIZE	DURATION	VISITS TO HOUSEHOLD	QUESTION-NAIRE SIZE	COST (US$M)	Time Series	Cross-sectional	Counter factual
POPULATION CENSUS	Full coverage	3-6 months	1	4-8	15-25	X	√√	X
AGRICULTURAL CENSUS/ SURVEY	20 000-40 000	1-1.5 years	2-4	8-12	8-12	X	√√	X
LSMS/ INTEGRATED SURVEY	5 000-10 000	1-1.5 years	2	40+	1-2	X	√	√√
HOUSEHOLD BUDGET SURVEY	4 000-10 000	1-1.5 years	15-25	15-20	1-2	X	X	√√
COMMUNITY SURVEY	100-500	4-6 months	1	4-6	0.2-0.4	√	√	X
SERVICE DELIVERY SURVEY (CWIQ)	10 000-15 000	2-3 months	1	8	0.2-0.4	√√	√	X
FOCUS GROUP INTERVIEWS	40-50	2-3 months	1-3	-	0.05-0.1	√	X	√
WINDSCREEN SURVEY	10-20	2-3 weeks	0		0.01	√	X	X

Applying the tools for M&E analysis

Which tools are best for monitoring ARD programme results? There is, of course, no right answer to this question; it all depends on what one is trying to do. Box 13 compares each of the key characteristics for all of the above surveys. The numbers are indicative only, particularly the costs of the different types of survey, because it is not always easy to separate out investment costs, which includes the purchase and rehabilitation of vehicles, computers, etc. with recurrent costs. Nevertheless, they do help to highlight the differences between the various types of surveys.

The final three columns need explanation. When tracking programme results, the M&E analyst basically uses the data to make comparisons, which may be of three types:

- comparisons over time (time series analysis);

- comparisons over space (subnational comparisons);
- counterfactual comparisons (with/without project/programme).

Each of these tasks requires different tools. Two ticks signify that the tool is well-adapted to the task; one tick, that the tool is adequate; and one cross, that it is not suitable.

Comparisons over time

Essentially, such comparisons involve tracking one or more indicators over time to see how they change. The first use of this time series analysis is generally to provide short-term feedback to policy-makers and programme implementers to allow them to make adjustments to the programme during its implementation. The prerequisite for this task is a continuous and reliable supply of consistent data. Most probably, the information will be needed on an annual basis, likely at a fixed point in the year, some months before the budget preparation process is due to start. This therefore rules out some of the larger surveys, since they are most unlikely to be conducted more than once every three to five years. What is required is a simple set of core questions that are quick and easy to collect and process, and that will be collected repeatedly every year. A service delivery survey such as the Core Welfare Indicators Questionnaire (CWIQ) fits the bill.

It is important to ensure the consistency of methodology over time and a consistent and uninterrupted supply of data.

However, while the service delivery survey may be suitable for monitoring the access, use and satisfaction indicators, the problem remains of how to monitor the longer-term physical changes resulting from the various ARD programmes. What is needed in terms of data is, simply, consistent annual reporting on agricultural production, yields and areas.[3] The dilemma here is that these are priority indicators that everyone needs, yet few countries currently have the statistical capacity to generate the necessary information with sufficient accuracy and timeliness to satisfy this basic demand.

Baseline surveys

At this point, the issue should be raised of the **baseline** survey and the case made that, where statistical capacity is weak, acquiring the baseline data does not necessarily require a heavy-duty baseline survey. Baseline data are required for two purposes. First, they are needed to provide the programme designers (planners) and implementers (managers) with as accurate and detailed a picture of the current status of the population in the target area as possible. This information is used to identify the needs of the intended beneficiary groups and to orient the project

3 This should cover not only crop production, but also livestock, forestry and fisheries.

design toward satisfying them. These data are therefore needed *before* the start of the project or programme, during the project preparation phase. A multi-sectoral integrated household survey, such as the LSMS, is well-suited for this purpose, but it may not always be cost-effective to undertake one. Alternatively, it may be possible to assess and understand the needs of the region using more qualitative approaches, such as participant observations or focus group interviews. Even though they are not statistically representative, such instruments can provide rich insight into the concerns and priorities of the project/programme beneficiaries.

Baseline data are important, but may not require a large-scale baseline survey.

The second purpose of baseline data is to provide the initial values of indicators to be monitored throughout the life of the project or programme. It is very important that the initial readings for these indicators be taken as soon as possible, preferably before the project or programme becomes effective. This may not require a full-scale multi-topic baseline survey, and could just be the establishment of the monitoring mechanisms and the starting values for these indicators. Consequently, one should embark on a baseline survey with caution, as it can pull scarce resources away just when they are needed most for other critical tasks.

It is important to ensure that the baseline survey sample includes a control group of non-beneficiaries against which the project beneficiaries can be compared. This is particularly important in subsequent impact evaluation of the intervention and provides the basis for assessing "with" and "without" project impact in the targeted area.

Panel surveys

Another question that arises at this stage is "What about using **panel surveys**?" Up to now, mention has been made of repeating cross-sectional surveys – that is, drawing a new sample of households every year while keeping the questionnaire itself constant. This is the correct way of monitoring overall changes in poverty levels and living conditions, etc. But the panel survey is different: it keeps the same sample of households (the panel) over several years, and the panel members are re-interviewed each year. This is another way of tracking poverty, by observing who moves in and who moves out of poverty. It highlights and identifies issues and trends that cannot be captured using traditional sampling procedures, and which may merit further research or consideration. Powerful though this instrument is, however, it should be noted that the panel that was

Panel surveys are powerful but difficult analytical tools.

randomly selected in Year 1 to represent the population at that time will no longer be representative of the population in subsequent years. Therefore, it is not suitable for tracking changes in living standards at the aggregate level. It should also be noted that panel studies can be extremely complex to carry out, because households may be highly mobile and because the composition of the household itself changes from year to year. It may therefore be necessary to commission out such surveys to a university or research centre, which may be better placed to provide the level of dedicated supervision needed for complex studies of this type.

Comparisons over space

This involves making comparisons at the subnational level between different geographic areas, which are particularly relevant to ARD programmes. With the growing emphasis being placed on decentralized decision-making, there is need for disaggregated data that allow estimates and indicators to be produced at the district level or below. The constraint in this case is sample size. If one were to take a country with, for example, 100 districts, and apply the rule of thumb of 500 to 600 households per analytical domain, sample sizes of 50 000 to 60 000 households would be required. This is beyond the capabilities of most NSOs, and alternative avenues must be sought.

It is vital to think through the survey logistics before embarking on large sample surveys.

One option would be to use a rotating sample and cover, say, one-third of the districts each year. Thus, any one particular district would be covered once every three years. Another option would be to drop the idea of a centrally administered survey and to concentrate on building up capacity at the district level to undertake simple district level surveys. Over time, this may well be the best solution, but currently, it is highly questionable whether any of the less developed countries would have the capacity to undertake such survey work at the lower administrative levels. A third option would be to employ a combination of tools and to use them to impute values at highly disaggregated levels. These techniques have been successfully developed and used in the context of poverty mapping. They involve taking advantage of the *breadth* of coverage of population census data and the *depth* of coverage of a recent, integrated household survey, and using the two instruments to estimate poverty incidence variables at the level of the lowest administrative units. The fourth and possibly most promising option would be to de-emphasize the idea of collecting district-level information through probability-based household surveys and to focus instead on the analysis of administrative records, or to use community surveys to collect the data.

Counterfactual comparisons

These comparisons address such questions as "What would have happened had there been no project?" or "What if the project had been differently designed?" They open up opportunities for multi-sectoral and multidimensional modelling. Here, the analysis goes beyond the question of "Are agricultural incomes rising?" It probes the data to discover why they are or are not, and what they would have been like had there been no intervention. An integrated multi-topic survey is probably one of the best instruments to address such questions, but there are other approaches that can be used as well. Qualitative methods work well and provide insights that structured formal surveys only seldom do. Another option is to combine service delivery surveys with household budget surveys, which provide very nearly the same information base as the integrated surveys.

In conclusion, there are a number of tools now available for monitoring and evaluating ARD sector programmes, each with its own strengths and weaknesses. These need to be very carefully assessed because the collection and production of statistics data is not an inexpensive undertaking.

Strengthening National Statistical System capacity

One must be careful not to generalize too much, but in many countries, NSSs have been severely under-resourced and have been unable to deliver both in terms of timeliness and data reliability. Their primary responsibilities are to collect and be the custodian of the entire nation's official statistics. Yet, the national statistics databases suffer from gaps or are filled with imputed values that are themselves prone to gross errors. This has led users to become increasingly dismissive of the efforts of the NSO, and in the process to stop providing feedback on where and how the databases could be improved. The inevitable knock-on effect of this is that resources for statistics are further reduced. In Africa today, there is almost no NSO that is functioning without significant flows of donor funds. Yet, donor support has not been well coordinated and has actually had a distorting effect on survey programmes and priorities, leading to an unproductive and wasteful use of statistics services.

Agricultural and rural sector statistics cover a broad range of topics for many different primary products, including production, inputs, trade, resources, consumption and prices. The list becomes much broader, if one adds closely related topics such as the environment and climate statistics. They come from many different sources, both governmental and non-governmental. They may come from institutions operating within the agriculture and rural sector as well as from outside. Some come from international sources. The primary responsibility for collating all these data rests mainly either with the Ministry of Agriculture or with the NSO. Until the 1990s, most national statistical survey programmes consisted of traditional sectoral-focused surveys, including Labour Force Surveys (LFSs), health and education surveys and

Household Budget Surveys (HBSs), as well as agricultural surveys. For better-off countries, this continues to be the case, except that multi-topic household surveys have been added to the list. For the poorest countries, however, as resources became increasingly constrained, cuts and adjustments had to be made. Given the high cost of household surveys, the move towards integrated surveys was considered good value for the money, because multiple objectives could be met using just the one survey instrument. In these countries, multi-subject surveys started to replace other household surveys. While this has a number of advantages, the production of agricultural statistics has suffered in the process, because agricultural surveys – traditionally used to collect information on production, area, yield and prices – have been conducted with increasingly less frequency.

When agricultural surveys are carried out by Ministries of Agriculture, they often use an area-based sample frame and take the holding as the basic unit of enumeration. When carried out by the NSO, it is more likely that they will be integrated into the household survey programme and use a population-based frame with the household serving as the unit of enumeration. While this is perfectly satisfactory for the analysis of the many dimensions of household living standards, it is a less efficient design for agricultural statistics. The trend towards integration has meant that, in a number of poor countries, independent agricultural surveys have almost ceased to be conducted. Instead, an agricultural module has been added to an integrated programme of household surveys. Again, from the point of view of agricultural data, this has required compromises that have reduced the quality of the core agricultural data.[4]

Budget cuts have also meant that NSOs have had to lay off staff. One of the primary assets that many of them had built up was a permanent cadre of field staff spread across the country and living frequently in or near the actual primary sampling units of an NSO master sample frame. They were trained and ready to conduct any survey to which they might be assigned. This gave the NSO an enormous comparative advantage over other agencies. But with the layoffs, this advantage has been lost. In many cases, the permanent staff have been replaced with mobile teams of enumerators – again, cost-effective but statistically less satisfactory, because of language problems in the different regions and because any outsider arriving in the village was always treated with more suspicion than a permanent enumerator.

In reviewing the performance of NSOs over the past decade, one might conclude that when it comes to the basic task of survey implementation, NSOs still have a significant comparative advantage over other agencies. Their capacity for analysis is weak, however, and they are mostly not appropriately structured

4 For instance, when collecting standard household information, particularly information on incomes and expenditures, the reference periods are linked to the standard calendar month or week. For agricultural statistics, however, the more logical reference period is the agricultural season – but the schedule of visits to the household in an integrated survey tend to ignore this for operational reasons.

to take on the deeper analysis and exploitation of the surveys. In particular, NSOs with weak capacity should be wary of undertaking quasi-experimental surveys, or panel surveys requiring a high level of supervisory competence, if there is any danger that these may negatively affect their ability to deliver their core statistics programme. New alliances need to be formed with universities and research centres so that there would be a greater sharing and pooling of data gathering and surveying expertise.

The issue of *data access* remains a major issue for many countries. NSOs are extremely guarded about granting access to the primary data sets claiming in many cases that this would be a breach of confidentiality. The real reasons may be more related to a lack of technical capacity, particularly in the areas of data archiving and storage; unwillingness of management to allocate sufficient resources to build up competencies in this area; and fear of political interference.

Impact of devolution and decentralization

Any discussion on the evolving role of M&E and how it can be supported by the NSS needs to make reference to the challenge presented by the growing trend towards devolution and decentralization, and the parallel growth in demand for subnational (district-level) statistics. Subnational issues have become increasingly important in many countries. This interest parallels the increase in fiscal responsibilities of subnational governments and the evolving trend toward decentralization. Many countries now pursue broader decentralization reforms for a number of political and economic reasons, as well as for poverty reduction. Decentralized decision-making can bring governments closer to the people, overcome information asymmetries, and enhance transparency and accountability. While the arguments for pursuing a programme of decentralization are persuasive, its implementation is not easy. In many countries, the technical capacity of government departments at the subnational level is extremely weak, thus requiring a major capacity-building programme in all areas. This includes local-level capacity building in programme planning, implementation and M&E. With reference to M&E in particular, the relationship between central and subnational systems is complex, since subnational M&E systems have to respond to subnational needs as well as contribute to national needs, and the requirements of each are not necessarily the same. Essentially, the data are needed at much lower levels of disaggregation. Ideally, the goal would be to have results available at the level of the lowest administrative unit – the village or parish – and to make the results available to the communities themselves so that they can compare their village against other villages in their district, and their district against other districts in the country.

But the primary responsibility of the NSO is to provide reliable and timely statistics information at the national level, and its ability to do this may be jeopardized if it tries to spread its slender resources too thinly. If the NSO or other agencies within the NSS were simply to expand the coverage of

their ongoing surveys, the sample sizes would be prohibitively large – almost certainly beyond the resources of the NSO. Alternative solutions have to be sought. To begin with, the role of the NSO almost certainly has to change from survey implementation to training and quality assurance. Additional suggestions include: using local field resources (enumerators); using a rotating sample so that not all districts are covered at once; or conducting community-level surveys rather than household-level surveys. These and other options were discussed earlier in the chapter.

Chapter 4
The institutional framework

When it comes to the M&E of sectoral programmes and national development and poverty reduction strategies, a large number of different institutions become involved, and problems of coordination and programme management become major issues. This involves not only horizontal collaboration across different sectors, but also the creation and strengthening of vertical ties linking communities and local governments to central authorities, and linking national governments to international agencies. The final challenge for building up monitoring and evaluation competencies is neither technical nor conceptual, but lies in ensuring that the required incentive structure and institutional capacity is created to be able to perform these functions. The challenge is particularly daunting in that the countries that are the poorest and that most urgently need viable poverty monitoring systems are also those where statistical and analytical capacity is weakest and poverty monitoring resources are most limited. The discussion begins by recognizing that important changes are taking place with respect to the strengthening both M&E capacity and the statistical infrastructure, but that there is insufficient interaction between these two communities of practice despite the obvious synergies.

The M&E framework

An important part of the preparation of this Sourcebook has been the field validation in five countries (Cambodia, Nicaragua, Nigeria, Senegal and the United Republic of Tanzania) of the indicators and M&E methodology that it advocates. In each country, a consultant was recruited to undertake an overall

Box 14. How do we know if a Poverty Reduction Strategy is effective?

First, a poverty monitoring system is needed to track key indicators over time and space, and to see if they change as a result of the strategy. Countries must be able to set up a poverty monitoring system in order to define key indicators, track them over time, and see what changes have taken place. Many countries already have poverty monitoring systems in place, so the task is to assess their adequacy and strengthen them as necessary. Experience shows that elements such as the tracking of public expenditures and outputs, and quick monitoring of household well-being need special attention. Also, participatory data collection methods and qualitative information give a different perspective and should not be overlooked.

Second, rigorous evaluations should be done selectively to assess the impact on poverty of interventions that are key components of the strategy. Countries must decide when it makes sense to do a rigorous impact evaluation, and how to design and carry it out, including what data are needed for different methodologies and how to obtain the data.

Other types of evaluation, such as assessing the process of formulating a poverty reduction strategy, can also be useful. Another challenging issue is how to evaluate the impact of poverty reduction strategies as a whole, as opposed to the impact of specific components of a strategy such as programmes or single policies. The key point made here is that a solid monitoring system will provide the basic data necessary to conduct such evaluations, should the need arise in the future.

Both monitoring and evaluation activities need to be carried out by competent institutions that have strong links to key decision-makers if they are to be useful in the design and implementation of a poverty reduction strategy. Much monitoring and evaluation takes place without adequate development of in-country capacity and without strong links to key decision-making processes; thus, precious opportunities to learn what works and what does not are lost. Countries need to build capacity and, in particular, strengthen the processes that provide policy-makers and others with feedback on the impact of policies and programs.

Dissemination of results is critical for use. Results that are not widely disseminated through mechanisms tailored to different groups in civil society will not be used, and the resources that were spent in getting such results will be wasted.

Non-governmental actors – research institutions, civil society organizations, special-interest and advocacy groups and others – have an important role to play in the design of the monitoring and evaluation system, in carrying out monitoring and evaluation activities, and in using the results.

World Bank, 2001, PRSP Sourcebook

assessment of current practices and to compare them with what is proposed in the Sourcebook. The exercise culminated in national workshops in each country, in which national participants were given the opportunity to present the different aspects of their own national monitoring and evaluation activities and to compare them with the recommendations in the early draft of the Sourcebook. The deliberations of the workshops have significantly enriched the final Sourcebook, and most of the boxes that appear in this chapter have been extracted from the workshop summaries.

Most countries already support numerous ongoing M&E activities. The challenge is to coordinate the different programmes cross sectorally.

Box 14 is taken from the World Bank poverty website. Not only does it illustrate the wide range of activities that need to be undertaken, but more importantly, the large number of disparate institutions that need to be involved. Whether countries already have active ongoing national M&E programmes, or whether they are starting from scratch, those embarking on a PRS usually include, during the preparatory phase, a full review of ongoing M&E activities at all levels – project, sector, national – and an assessment of their capacity-building requirements. It would be rare to undertake such a review and not discover a large number of formal or informal M&E activities already taking place. In fact, the situation may appear chaotic and disorganized. This should not be a deterrent and should certainly not be a reason for trying to disband or reject such initiatives. The goal should be one of inclusion, not exclusion, and of creating a network of M&E units; Cambodia provides a good example (Box 15).

In some countries, the relationship between the different network members is formal and hierarchical; in others, it is much looser. One of the main reasons for establishing a network is to encourage knowledge sharing and the adoption of common reporting standards, so that data from different projects and programmes can be aggregated or compared.

Most programmes with an M&E component will have an M&E officer or unit, or possibly share one. The PRS is no exception. The PRS M&E unit may be located anywhere in the government system – or even outside it. There may be competition among potentially eligible institutions wanting to house the unit as resources are likely to come with it. In many cases, such a unit will be attached directly to the Ministry or body responsible for overseeing the overall implementation of the PRS. In some cases, the national M&E unit and the Poverty Monitoring Unit have been merged into one; in others, they have remained separate but linked. The United Republic of Tanzania provides a particularly good example of an integrated system bringing together what had previously been a number of disparate and separate monitoring activities (Box 16).

Box 15. The M&E system of Cambodia's Ministry of Agriculture, Forestry and Fisheries

As part of its Public Financial Management Reform (PFM), the Ministry of Economy and Finance (MEF) has chosen the Ministry of Agriculture, Forestry and Fisheries (MAFF) as the pilot line ministry to introduce and demonstrate the application of the Ministry Strategic Budget Framework (MSBF) through an efficient and effective delivery of services.

Individual programmes and sub-programmes need to be monitored so that resources are allocated based on performance. This requires a well-functioning monitoring and evaluation (M&E) system that regularly collects information from individual activities and assesses their contributions to meeting the Ministries' strategic goals.

The M&E system for programme budgeting relies on the programme structure described in the MSBF. MAFF's resources are assigned to a three-tiered structure of programmes, sub-programmes and activities. Each programme can have any number of sub-programmes and activities. The MAFF M&E system is built around a results chain with a small number of carefully selected indicators to be monitored at each level, as follows:

TYPE OF INDICATOR	WHAT IS MEASURED	INDICATOR	NO. OF INDICATORS
Goal (programme)	Results from the combined effect of a multiple outcome toward a development condition at the programme level.	Use of outcomes and sustained positive development change.	3
Outcome (sub-programme)	Results from the outputs generated by multiple activities, projects and partners.	Use of outputs and sustained production of benefits.	3
Output (activity)	The good or service that is produced through work performed in activities.	The output produced by the activity, expressed as a measurable indicator.	1 indicator per output

The M&E unit is at the centre of all M&E activities. At the project level, it would most likely appear on the organizational chart near the project manager, and the M&E officer heading the unit would be part of the management team. At the sector level, the unit may be located in the Ministry and closely associated with the planning department. At the PRS level, the M&E unit will be close to the PRS oversight committee (or equivalent); it may even serve as the secretariat to the committee.

Box 16. The Poverty Reduction Strategy Monitoring Master Plan (MUKUKUTA) of the United Republic of Tanzania

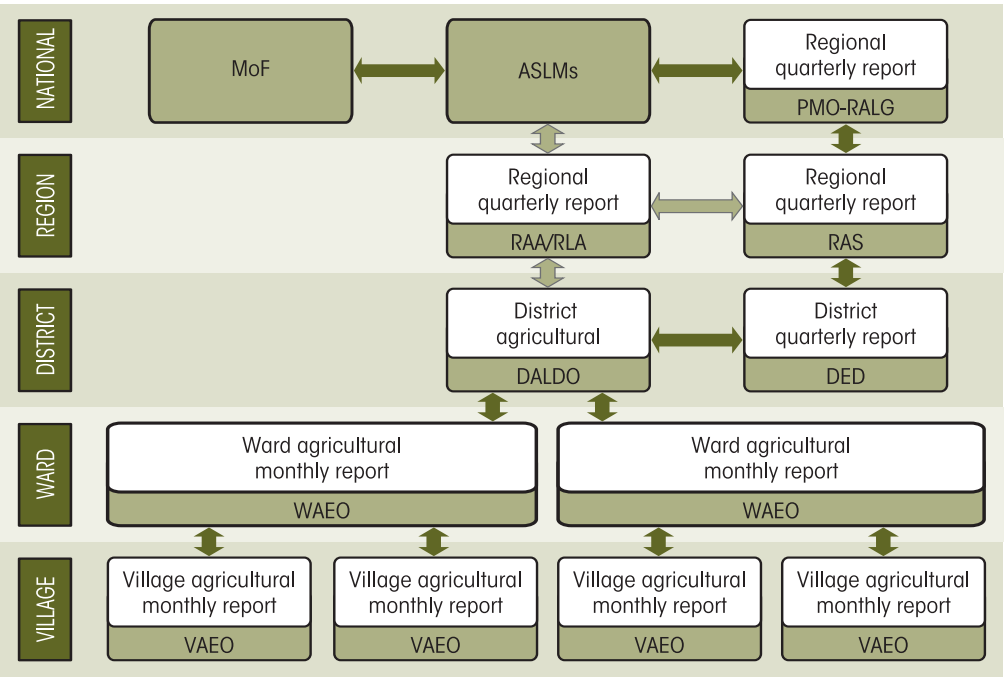
M&E in Tanzania is done at different levels of government and the overall framework is coordinated by the Ministry of Planning, Economy and Empowerment (MPEE). At the national (macro) level, information is obtained from a wide range of institutions including ministries, departments and agencies (MDAs) and local government authorities (LGAs), which have Management Information Systems (MIS) and performance reporting requirements linked to their Strategic Plans and Budgets.

Early results from sector plans monitored through subnational (sector) and national-level indicators provide hints to the government on what interventions are needed to improve the sector's performance in relation to MUKUKUTA targets.

Use of M&E results as basis for budget allocation

The MUKUKUTA Monitoring System provides an integrated approach to output and outcome reporting within Government, and provides analysis of changes in relation to goals and operational targets of MUKUKUTA. These then inform decisions about national planning, budgeting and public expenditure management.

Planning processes begin with development goals as articulated in the Vision 2025. In MUKUKUTA, these goals are translated into operational targets and are linked to cluster strategies, which provide the national medium-term framework for planning. The Strategic Plans of each MDAs and LGAs translate MUKUKUTA into budgets and action plans (programmes, targets and activities).



Box 17. M&E Technical Committee – sample Terms of Reference

In most countries, the head of the M&E Unit also chairs an M&E technical committee, comprising representatives of the different network nodes – the heads of other sectoral M&E units – and other interested and involved stakeholders, both from within and outside government. The National Statistics Office (NSO) should be a core member of the coordinating committee. The relationship between the M&E Unit, which essentially heads the national M&E network, and the NSO, which heads the National Statistics System (NSS), is a critical one, and not always easy as a result of occasional conflicting priorities. The main responsibilities of an M&E Technical Committee may include:

- defining, and ultimately delivering a national M&E Action Plan;
- agreeing on and ensuring adherence to national standards, definitions and methodologies;
- facilitating the smooth flow of timely information between the various members.

Where an M&E Study Fund has been set up to finance technical studies, workshops and other knowledge-sharing events, the M&E Unit shall have the responsibility for managing the fund, but the Committee shall have the responsibility for approving the studies that it will finance.

The M&E Unit is responsible for producing timely reports and will accordingly maintain a large database of indicators. This database will regularly need updating and be used for the preparation of the reports. The Unit will also be responsible for commissioning studies and evaluations when needed.

The head of the unit, the M&E officer, needs excellent skills in communication and in coordinating and bringing people together. There is good evidence that the best examples of successful M&E programmes are to be found where the head of the unit plays the role of the M&E advocate with conviction and passion. The position should clearly be a senior one as it requires a combination of good analytical skills and good communication skills. The office must be able to understand the information needs of management and of other stakeholders – he or she will be listened to at the highest levels.

The functions of the M&E unit are described in Box 17. They include the preparation of regular monitoring reports on progress and achievements, as well as the commissioning of a wide range of evaluation studies on different aspects of the PRS. This necessarily involves consolidating the various sector reports prepared by the sector M&E units. The relationship between the central M&E unit

Box 18. National Planning, Monitoring and Evaluation (PM&E) Workshop in Nigeria

The annual National Planning, Monitoring and Evaluation (PM&E) Workshop is a special feature of the M&E system in Nigeria. The Workshop provides a forum where all the key professionals in the M&E system as well as those interested in the M&E results meet to discuss and review progress in implementation of development projects in the country.

The main objective of the workshop is to bring together the PM&E officials in the state Agricultural Development Projects (ADPs) and other national programmes to discuss the issues relating to efficiency and effectiveness of the M&E system in the country. In addition to reviewing progress on project implementation, the forum also serves as an occasion to build capacity of M&E professionals in the country. The Workshop is also an instrument for assessing and reviewing the achievement of stated government policy objectives, targets for agriculture and rural development (ARD) programmes as well as the functioning of M&E in the country.

This annual meeting of M&E professionals started in the late 1970s with the establishment of the World Bank-assisted ADPs in Nigeria. Initially, it was known as the National M&E Seminar, and participation was led by the then Agricultural Project Monitoring and Evaluation Unit (APMEU) in the Federal Ministry of Agriculture and Water Resources. After the merger of the Federal Agricultural Coordinating Unit (FACU) with APMEU in 2001 to form a Project Coordinating Unit (PCU), the Seminar was renamed the National Planning, Monitoring and Evaluation Workshop, and its participation was extended. Currently, the Project Coordinating Unit (PCU) takes the lead in organizing and coordinating the activities related to the Workshop. The Workshop is hosted by the states on a rotational basis but it invariably receives representations from other leading national institutions involved in M&E, including:

- Central Bank of Nigeria (CBN)
- the National Bureau of Statistics (NBS)
- the National Planning Commission (NPC)

The Workshop receives the patronage of political heads from the Federal Ministry of Agriculture and Water Resources and the host state, who deliver the opening addresses. Efforts are also made to seek the participation of donors and development partners in the Workshop. Goodwill messages from country leaders of the donor community are a common feature in the Opening Session. The Plenary Session entails presentations and discussions of invited technical papers by renowned scholars, from within the M&E system as well as in academia, on topical issues relating to PM&E development. This follows the presentation of reports by the state ADPs and other agencies on their PM&E activities during the preceding year and the Action Plan for the next year. The reports are thoroughly discussed and the necessary resolutions are passed. At the end of the Workshop, a communiqué is issued. The Proceedings of the Workshop are later sent to relevant authorities for necessary follow-up actions on the decision taken in the Workshop.

and sector units varies enormously. The goal for countries is to establish an all-government M&E system with the central unit at its head and with each of the sector M&E units responsible for sector-level reporting. In principle, coordination is managed by creating a national M&E technical committee chaired by the head of the M&E unit (Box 17). Clearly, this implies a degree of authority over the sector units. The reality on the ground may be less clear. In many cases, sector and project M&E units continue to operate with considerable autonomy in parallel with, and independently of, the PRS central unit. One of the more important functions of the unit is that of advocacy, promoting the concept of management by results, organizing workshops to review the outcomes of various monitoring activities, and discussing lessons learned to be drawn from them. In Nigeria, where there is a wide range of M&E initiatives operating at different levels, an M&E workshop is convened annually to bring the various M&E practitioners together (Box 18).

The statistics framework

In parallel with the growth of interest in the monitoring and evaluation of national development programmes, there has been similar interest in the rehabilitation of the NSS. The NSS comprises all the institutions and agencies that contribute in some way to the national statistics databank. This includes line ministries, Customs and Excise, the Central Bank and others. The apex institution for the NSS is the NSO. In effect, the NSS is the national statistics network – equivalent to the M&E network described earlier. Many of these institutions are the same as those represented on the M&E technical committee, but there is no guarantee that their representatives will be the same as those represented on the NSS. Thus, one may find two communities of practice within one country, the M&E community and the statistics community. Both work on parallel issues, but not necessarily communicating or working together, except possibly at the highest level.

The question may be asked "What is the difference between M&E and statistics?" It is hoped that readers of the Sourcebook should by now have a clearer understanding of the different natures of the two entities, but even so, it can still be difficult to distinguish the two from each other. Box 19 illustrates how Nicaragua has confronted the challenge. What is clear is that, although they have evolved separately and a have different mandates, there are still large areas of common ground where their activities overlap and where there is great potential for working together for mutual benefit. The monitoring of ARD programmes and the PRS generates a constant stream of demands.

In general, the priority indicators and the basic agricultural and rural statistics needed for monitoring ARD programmes, described in the previous chapters, are the same core statistics that the NSSs should be generating, except that few NSOs currently include service delivery monitoring in their core survey programme. However, given the fact that such data are relevant not just to monitoring ARD programmes, but also for monitoring service delivery across other sectors, NSOs

Box 19. Nicaragua – Linking the M&E activities more closely with the National Statistical System

Nicaragua is currently upgrading its statistical services. It is also keen to strengthen its monitoring and evaluation capabilities with a view to improving the quality of public enterprise management. In many countries, there is a significant gap between what information is desired for M&E purposes and what is being provided by the NSS; Nicaragua is no exception. In the course of reviewing its needs, both in the area of statistics and M&E, it has become clear that, despite a number of areas of overlap, there has been relatively little communication or collaboration between the statisticians, on the one hand, and the M&E practitioners, on the other. Statistical priorities have traditionally been largely determined within the statistical system itself, and M&E systems have been set up without seeking a technical input from the offices of the NSS. It is generally agreed that improved coordination would benefit everyone and would allow for much more efficient use of national resources.

A number of steps are being taken to rectify the situation. The most important has been the introduction of a new National Strategy for Statistical Development (ENDE), in which a number of sectoral forums are being established to ensure that sectoral information needs are fully addressed. The Forum for Agricultural Development in particular will be very active in reviewing the statistical work programme and ensuring that it is capable of providing at least a proportion of the most urgently needed statistics for monitoring and evaluation. At the same time, the position of the officer responsible for the M&E system has been upgraded to a higher level. The aim is to raise the level of advocacy for M&E and to make sure that the needs of the M&E system are recognized by the NSS and given appropriate attention.

should be receptive to this request. In the end, it comes down to negotiation. The additional burden to the NSS need not be excessive, but at the time of the negotiation, it is important that a timetable be specified for when the results will be needed, and with what frequency a survey would need to be repeated. It is not a one-sided negotiation: in most countries, there is no stipulation that the NSO has to be the sole agency used to supply the data. It is also a competitive open market situation, and other public or private sector institutions may be capable of doing the job better and/or cheaper.

The first responsibility of the NSO is that of serving as the chief compiler and custodian of all official national statistics. This is its primary mandated

Box 20. Senegal's Reformed National Statistical System

The Senegalese National Statistical System has the following vision: "To become a robust System wich is well coordinated and responsive to users' needs". The ongoing reforms will be implemented over a medium- to long-term time frame to ensure that all actors are on board and that their roles are correctly understood. The vision will be built on four key pillars:

- Strengthening the institutional framework
- Improvement of the quality of statistical products
- Dissemination and promotion of the use of statistics, analysis and research
- Strengthening capacity for an effective statistical system

The reformed statistical system is being built around the values of transparency; feasibility; efficiency and adaptability. The overall work programme will be shaped by the needs of the users and will ensure that international commitments are honoured.

The lead institution is the National Agency for Statistics and Demography (NASD). NASD has been granted a large degree of autonomy and will be a reference centre with resources in line with the magnitude of its responsibilities and duties. The NASD is supervised by the following authorities: the National Council of Statistics, which approves the Annual National Programme of Statistical Activities, and the Technical Committee of Statistical Programmes in charge of the preparation documentation to be submitted for approval by the National Council of Statistics. The Technical Committee also oversees the implementation of the decisions of the National Council.

responsibility. The NSO is under pressure from a wide range of users competing for scarce statistics information. It will try to balance the different demands. Further, one expects it to put the provision of statistical support for the monitoring and evaluation of national development programmes high on the priority list, but the demands for M&E data could occasionally conflict with other demands and may not always be given the highest level of protection, certainly not unless the request comes with extra resources.

Both monitoring and evaluation have been given a significant boost with the growth in popularity of the concept of management by results. Evidence-based development requires underpinning by statistical information and data. A second boost was provided by the MDGs and by the PRS, both highlighting poverty reduction as the prime goal for all development efforts. Evidence must

be provided that poverty is indeed falling, and must be supplied through the NSS. The most significant implication of this growth in demand comes from the fact that the demand is increasingly "home-grown" – it comes from within the country, rather than from the donors outside. Without such a growth of domestic demand, it is difficult to see how any strengthening of the statistical infrastructure could possibly be sustainable.

In addition to this growth in domestic demand, there has been an evident movement by the donor community to jointly commit to supporting the strengthening of NSSs, and in a coordinated manner. In order to be eligible for international support, it will first be necessary for the national office to prepare a strategy for strengthening the NSS. The undertaking of a major overhaul of the NSS is not a necessary condition for establishing an M&E capability in the country, but for many countries where the statistical infrastructure is weak, it is strongly advised that, at the very least, a review of ARD statistics be carried out. Senegal is one country currently reviewing its statistics system with a view to creating a more autonomous and effective NSS (Box 20).

The international framework

In conclusion, the challenges of M&E of ARD programmes need also to be addressed at the international level. The universal acceptance of the MDGs represents a global commitment to lift the poorest of the poor out of poverty. It establishes a demand for M&E at the very highest level. It will be necessary to report in 2015 on whether or not the goals have been achieved. Importantly, well before then, the mechanisms must be set up to track progress towards their achievement, and stakeholders alerted to issues of concern where countries or regions are clearly off-track – and in a timely manner so that corrective action can be taken. To achieve the MDGs, the international community must assist more than one billion people out of extreme poverty. Of these, 70 percent live in rural areas and depend on agriculture for their livelihood. The challenge is to understand how, where and when agriculture can make the greatest contribution to achieving the MDGs. Even though ARD do not have a specific MDG, they do make a major contribution towards two of them, MDG 1 and 2, and reinforce or contribute to at least five others (Box 21).

Monitoring of the MGDs is managed globally by the United Nations system, including the World Bank and IMF. The specialized agencies are responsible for compiling the indicators relevant to their particular sector. With respect to the monitoring of ARD, the relevant agency is FAO. The Organization does not collect its own primary data, but is essentially a source of secondary data; it compiles and distils data from a range of different primary sources, mainly directly from member countries, but also from global satellite networks. For country reporting, use is generally made of indicators compiled from national sources, generally by the NSS.

The process of compilation is complicated by the fact that data submitted by the country statistics offices are of extremely variable quality or are frequently

Box 21. Agriculture and the Millennium Development Goals

Progress in agriculture makes direct substantial contributions to:
Goal 1: Eradicate extreme poverty and hunger.
Goal 3: Promote gender equality and empower women.

Progress in agriculture reinforces two goals:
Goal 7: Ensure environmental sustainability.
Goal 8: Develop a global partnership for development
and these goals reinforce progress in agriculture.

Progress in agriculture makes indirect but vital contributions to:
Goal 2: Achieve universal primary education.
Goal 4: Reduce child mortality.
Goal 5: Improve maternal health.
Goal 6: Combat HIV/AIDS, malaria and other diseases.

Based on World Bank, 2005a

missing. A number of advanced techniques may be used to fill data gaps and provide a conceptual coherence that appears convincing at an international level. Yet, if gaps are too large or too many, their application becomes increasingly unsatisfactory. There is also the problem that different countries will have used different methodologies or definitions in computing a standard indicator. This, again, can be handled as long as the data submitted from the countries include full supporting metadata comprising the definitions and methodology used, sample size and known or anticipated biases. While each host agency may carry out significant transformations of the data to ensure standardization across countries, all of them are highly dependent on the outputs generated by the NSS. The relationship between these national and international institutions engaged in monitoring is not hierarchical, but complex and symbiotic, with the international institutions needing the outputs from the national institutions and vice versa. Ultimately, the global M&E network is only as strong as its weakest link. The donors have a vested interest in seeing that the capacity of national institutions is strengthened, if for no other reason than to maintain the standard of international reporting systems.

The role of development partners

The donor community has been indisputably among the strongest advocates for establishing good M&E procedures and for building up M&E capabilities. Donors have also provided strong support to the strengthening of national statistics capacity. Recent initiatives include the Marrakesh Action Plan for Statistics (MAPS). This plan, to which all donors have subscribed, is a measure of the commitment to support statistical capacity building in a coordinated manner. In order to receive the benefits of such support, countries are encouraged to establish their own priorities for statistical development through the preparation and implementation of National Statistical Development Strategies (NSDS).

The development of an NSDS is seen as the first step towards the major rehabilitation of the NSS. It provides a vision as to where the NSS should be in five to ten years and sets milestones for getting there. It also provides a framework for mobilizing, harnessing and leveraging resources, both national and international. An important guiding principle is that the NSDS should support the NSS as a whole, not just the NSO. Guidelines on how to undertake an NSDS have been prepared by Partnership in Statistics for Development in the 21st Century (PARIS21).

A five-step approach is proposed:

- Launch the process (NSDS Design Road Map).
- Assess the current status of the NSS.
- Develop the vision and identify strategic options.
- Prepare the implementation plan.
- Monitor the implementation plan.

Another important group of stakeholders within the international community is the international organizations, who are themselves responsible for maintaining databases for monitoring at the global level. These include the international finance agencies, the United Nations specialized agencies and the United Nations Statistics Department. With respect to ARD, the agency most concerned is FAO. FAO is mandated with the primary and unique international responsibility to produce statistics on agriculture, land, water, forests and aquaculture. FAO maintains the largest statistics data set on food and agriculture in the world. The Organization compiles and extracts data from a range of different primary sources, mainly from member countries, but also from global satellite networks. Responsible agencies in the countries include NSOs and Ministries of Agriculture. Where national capacity is weak, FAO can, in principle, supply countries with the requisite technical assistance.

Box 22. National Statistical Development Strategy essentials

The NSDS should be integrated into national development policy processes, taking into account regional and international commitments. It should:

- have political support and commitment, and be championed by high-level national official(s);
- be demand-focused and user-friendly, responding to needs and priorities for information to enable national governments to manage for results;
- develop statistics as a public good, funded from government budgets and complemented (where appropriate) by international support;
- be mainstreamed as part of national development policy, including for the design, monitoring and evaluation of Poverty Reduction Strategies, sector strategies, and other national development plans, as well as assessing progress toward the MDGs;
- respect all relevant legislation or regulation, recommending changes where appropriate;
- work within the national context, both cultural and institutional.

The NSDS should be developed in an inclusive way, incorporating results-based management principles and meet quality standards. It should:

- be the output of a consensus-building/advocacy process, which helps build commitment and partnerships, with clear processes for consultation throughout;
- be the output of genuinely nationally led, owned and inclusive participatory processes including all stakeholder groups (e.g. users, analysts, producers; government, private sector, civil society; international and regional organizations, bilateral donors and specialized agencies);
- incorporate results-based management principles in the design of the NSDS and manage its implementation with performance indicators (e.g. for the supply of statistical information, value for money, user satisfaction, governance, support to national policies, confidentiality) and a performance reporting, monitoring and evaluation plan;
- follow the values and principles portrayed by the United Nations Fundamental Principles of Official Statistics to produce useful high-quality data that will have the confidence of users of statistics;
- draw on international standards, recommendations and experience to capitalize on worldwide knowledge and for consistency between countries.

continue

The NSDS should be comprehensive and coherent and provide the basis for the sustainable development of statistics with quality (i.e. "fit for purpose"). It should:

- provide an assessment of the current status of the NSS (where we are), incorporating a comprehensive appraisal of statistical outputs measured against agreed criteria;
- maintain statistical production and procedures, building on existing activities and ongoing processes, during the design and implementation of the NSDS;
- provide a vision for national statistics (where we want to go), strategies to deliver the vision (how do we want to get there), which address institutional and organizational constraints and integrate all statistical planning frameworks, and performance indicators (how do we know we have arrived): it is not just a work plan;
- incorporate substrategies for leadership and management, financial management, human resources, communications, infrastructure (e.g. information technologies) and dissemination as well as the technical work areas (e.g. national accounts, poverty statistics, health statistics);
- set out an integrated statistical capacity building programme, which:
  - builds capacity to implement and adapt the strategy;
  - turns statistics into information through analysis, dissemination, publicity and user education;
  - is prioritized and timetabled (not everything can be done at once);
  - provides the framework for (annual) implementation work plans;
  - is realistic, pragmatic and flexible enough to cope with changes in priorities, new information needs and lessons learnt and is as easy to accomplish as possible;
- outline the financing requirements: responding to user needs but realistic about resources (implies prioritization, sequencing, cost effectiveness: e.g. considers alternative ways of compiling data such as administrative sources and sample surveys).

Chapter 5
Setting up an M&E strategy in agriculture and rural development

Choosing the right indicators is critical, but M&E is much more than simply selecting a set of pertinent indicators; it also involves the identification and strengthening of data systems to ensure that indicators can be captured in a timely and reliable fashion. A number of different institutions are likely to be involved, and institutional capacity has to be reviewed and, if necessary, strengthened. Above all, the internal demand for M&E has to be nurtured and promoted, and the concepts of management by results need to be progressively introduced at all levels. This is not a trivial exercise and is best undertaken by following a carefully sequenced action plan. The objective of the plan should be to improve the flow and use of indicators and other statistics for monitoring and evaluating ARD projects and programmes. Wherever possible, the action plan should be formulated within the framework of the PRS, or equivalent national development plan. The challenge is greatest in countries where conditions are less than ideal, that is, where demand is weak, evidence is not used to inform decision-making, and the stock and flow of information are irregular, unreliable and/or available with an unacceptable time lag.The first step is to undertake an assessment of current capacity. In some countries, the capacity may already be strong; in others, particularly the poorest or those that are in or just coming out of a conflict situation, the basic infrastructure may not be available at all. All countries stand to gain from this exercise, but the latter stand to gain the most. A key objective of the strategy is to help countries to map out a route that is most appropriate to their specific situation – and to monitor progress as they proceed along that path.

Countries should define a strategy for developing national M&E capacity as an integral part of their overall ARD strategy.

Countries should develop a national M&E capacity as an integral part of their overall ARD strategy. The first step is to undertake an assessment of current capacity. In some countries, the capacity may already be strong; in others, particularly the poorest or those that are in or just coming out of a conflict situation, there may be no basic infrastructure at all. All countries stand to gain from this exercise, but the latter stand to gain the most. A key objective of the strategy is to help countries to map out the route that is most appropriate to their specific situation – and to monitor progress as they proceed along that path.

To help carry out this exercise, the reader is specifically referred to the publications of Paris 21, in particular, *A guide to designing a national strategy for the development of statistics* (OCED/DCD, 2007).

The path to action consists of six steps:

1. Assessment and diagnosis
2. Review of indicators
3. Review of current data, sources and gaps
4. Development of action plans
5. Review of resource requirements
6. Monitoring the performance of the M&E action plan

Step 1: Assessment and diagnosis

The starting point is an assessment and diagnosis of the current situation. The assessment should recognize the M&E systems in operation and related initiatives, and build on them – not try to replace them. The purpose of the assessment should not be just to document the current state of M&E, but also to highlight and document what is and what is not working, and to assess the demand and interest for promoting a greater degree of results-oriented management. The cultivation of this demand must be a continuous and ongoing process, and is essential if the initiative is to move forward. If a strong advocate can be identified to take the lead in this work, chances of success will be significantly increased.

In order to facilitate the assessment process, a simple assessment survey is described in Annex 3. It includes a checklist of questions to be addressed. The checklist may be used in one of two ways. The short method is only suitable as a workshop exercise and is based on group discussions. It is appropriate if the primary objective is to raise awareness and stimulate interest in M&E capacity building in general. The full method is more suitable if the final objective is to prepare a proposal for an M&E capacity-building programme. Whichever the route used, the objective is to accumulate sufficient information to fill out a scorecard that will be used to rank the national M&E capacity on a scale of 1–100. The answers are obviously subjective; they can only be interpreted in general, not absolute terms. Countries scoring over 75 points would be considered to have strong overall capacity, and those with an overall score of less than 25 points would clearly have very limited capacity.

Step 2: Review of indicators

Step 2 is built around the analytical framework discussed in Chapter 2. Again, the starting point is to identify actual development actions, ongoing or planned, and to look at what indicators are currently being used. Then, for each of the development actions, an appropriate set of indicators is selected, using the methodology set out in Chapter 2 and the menu of indicators in Annex 1. This should be compared with the indicators currently being collected, and a

definitive list proposed. Each indicator should be accompanied by additional information regarding the source and periodicity required.

Step 3: Review of current data, sources and gaps

Step 3 then shifts the focus to the NSS, in particular the NSO, and the Ministry of Agriculture and other ARD Ministries who also contribute statistics to the system, and to compare what is available with what is needed – as identified in Step 2. This comparison aims to identify gaps in the data series and weaknesses in the data collection system that would need attention in order to meet these demands. The review does not just concern data; it must also consider the tools used to provide them. How, for instance, are production estimates obtained and with what frequency? Are there any alternative sources of information that can be used to check the official estimates? The institutions involved are also taken into consideration, including an assessment of their capacity to collect, process, and disseminate specified statistics information. The review should also ascertain whether there is any ongoing or planned programme of assistance to support the strengthening of the institution's capacity. Finally, it needs to include a review of the system itself, its management and the roles of the various stakeholders, thereby complementing and completing the work of Step 1.

Step 4: Development of action plans

Earlier chapters discuss the use of the logframe for developing a project by starting with a vision of the future (goal) and then conceptualizing a path to reach it. Step 4 is where that process begins. It is clearly important that there is a common or shared vision for the M&E system. Clearly, this will depend largely on the vision for the ARD strategy itself. Box 22 provides a useful summary of what an NSDS might include. A number of questions have to be addressed. Is the M&E system envisaged as a public service to be used to hold the management of public services accountable or, rather, as a tool allowing the beneficiaries themselves to be informed about M&E findings so that they can compare their situation with, for instance, that of their neighbouring district? And what about impact evaluation capacity? What capacity should be permanently available within the system and what could be contracted out?

Having defined the vision, how is it to be achieved? Will it be by strengthening what already exists or by putting something new in place? Will this be strictly an ARD M&E network or a component of a larger national M&E system? What are the priorities in terms of actions – to get some results as quickly as possible or to invest in staff training and capacity building first and then start to work on data provision?

Step 5: Review of resource requirements

Step 5 addresses the issue of the resources required. As part of the diagnostic in Step 3, an assessment should already have been made of the current costs of

Box 23. A results chain for building an M&E system

Result	Indicator/source
Economic growth and poverty reduction	NSS generates reliable, timely core indicators, e.g. • GDP per capita • % of children malnourished • Agricultural Production Index (plus other priority indicators)
IMPACT	
• M&E used for government decision-making, resource allocation, policy design • Parliament assesses and debates PRS performance • Media reports on M&E findings	User satisfaction survey to measure access, use and satisfaction with respect to M&E services
OUTCOMES	
• Established formal M&E framework or system, including reporting and feedback mechanisms • Revised statistics act • Established statistical databases and archiving • Annual agricultural survey • Annual service delivery survey	Annual review of work plan
OUTPUTS	
• Approved Action Plan • Approved NSDS • M&E and statistical training • Implemented Advocacy Programme • Training of analysts • Funding package secured	Annual review of work plan
INPUTS	

M&E and its outside financing. These figures need to be updated and the future costs of the system estimated. The final issue to be examined is the financing and how much might realistically be forthcoming in the form of international assistance. The national budget to foreign investment ratio should be calculated and projected over a period of, say, five to ten years. What is the amount of public funds currently being invested in M&E? What is the level of international support needed?

Step 6: Monitoring the performance of the M&E action plan

The final task is to define a system to monitor the performance of the M&E action plan itself. The system should identify what reports are to be submitted by whom and when, and should include indicators for each of the four levels – inputs, outputs, outcomes and impact – including details of how they are to be provided and with what regularity.

The end result should be to produce a development programme with a results chain that is very similar to the one shown in Box 23. The programme has been conceived using the same process described in Chapter 2, with which readers should now be familiar. The top of Box 12 shows that the intended long-term *impact* of a strengthened M&E capability is to contribute to the national development goals of economic growth and poverty reduction. The outcomes that will contribute to the achievement of these goals will be an increase in the range and number of users and in the overall level of satisfaction with the quality and relevance of the information database. The indicators will include standard early outcome indicators of access, use and satisfaction. At the start, the primary users or stakeholders may simply be those who have a financial or management interest in the project (donors, government). Later on, these should expand to include the beneficiaries, civil society at large, and their representatives in Parliament. Over time, one may also expect to see the media becoming more interested as well.

The expected changes in client behaviour depend on the ability of the M&E programme to generate useful outputs in terms of indicators, reports, studies and evaluations, workshops and training, etc. These changes must be assessed not just on the basis of the quantity produced, but also of the quality of the product. It is at this level where, in the first instance, capacity-building efforts are likely to be focused – particularly with respect to countries where conditions are less than ideal. Finally, at the bottom of the chain are the inputs that need to be made available in order to generate the outputs referred to above. Inputs include human resources, training workshops, equipment and financial resources, both national and international.

Bibliography

Bedi, T., Coudouel, A., Cox, M., Goldstein, M. & Thornton, N. 2005. *Beyond the numbers: understanding the institutions for monitoring poverty reduction strategies*. Washington, DC, USA, World Bank.

Booth, D. *Poverty monitoring systems: an analysis of institutional arrangements in Tanzania*. London, UK, Overseas Development Institute (ODI). Prepared for DFID and the World Bank.

Casley, D.J. & Kumar, K. 1987. *Project monitoring and evaluation in agriculture.* Washington, DC, USA, World Bank.

Casley, D.J. & Kumar, K. 1992. *The collection, analysis and use of monitoring and evaluation data*. Washington, DC, USA, World Bank.

DAC Network on Development Evaluation. 2002. *Glossary of key terms in evaluation*. Paris, France, OECD.

Deaton, A. 2002. *The analysis of household surveys: a microeconometric approach to development policy*. Washington, DC, USA, Poverty Reduction Group (PRMPR), World Bank.

FAO. 1982. *Estimation of crop areas and yields in agricultural statistics*. Economic and Social Development Paper. no. 22. Rome

FAO. 1988. World Conference on Agrarian Reform and Rural Development (WCARRD.):ten years of follow up. *Guidelines on socio-economic indicators for monitoring and evaluating agrarian reform and rural development*. Rome.

FAO. 1989. *Sampling methods for agricultural surveys*. FAO Statistical Development Series No. 3. Rome.

FAO.1996a. *A system of economic accounts for food and agriculture*. FAO Statistical Development Series No. 8. Rome.

FAO. 1996. *Multiple frame agricultural surveys. Vol. 1- Current surveys based on area and list sampling methods; Vol. 2 – Agricultural survey programmes based on area frame or dual frame (area and list) sample designs*. FAO Statistical Development Series No. 10. Rome.

FAO. 2005a. *A system of integrated agricultural census and survey*. Vol 1. *World Programme for the Census of Agriculture 2010*. FAO Statistical Development Series 11. Rome. Chapter 3.

FAO. 2005b. *Review of the state of world marine fishery resources*. FAO fisheries technical paper No. 457. Rome (available at ftp://ftp.fao.org/docrep/fao/007/y5852e/y5852e00.pdf).

IFAD/United Nations Administrative Committee on Coordination (ACC): Force on Rural Development Task. 2002. *Guiding principles of the design and use of monitoring and evaluation in rural development projects and programmes*. Rome.

Kusek, J. & Rist, R. 2004. *10 steps to a results-based monitoring and evaluation system: a Handbook for development practitioners*. Washington, DC, USA, World Bank.

Mackay, K. 1999. *Evaluation capacity development: a diagnostic guide and action framework*. ECD Working Paper Series. Washington, DC, USA. Operations Evaluation Department (OED), World Bank.

OECD. 2002. *Glossary of key terms in evaluation***.** Paris, France. Organisation for Economic Co-operation and Development (OECD).

OECD. 2004. *The Marrakech Action Plan for Statistics. Better data for better results. An action plan for improving development statistics.* Presented at the second international roundtable on managing for development results, Marrakech, Morocco, 4–5 February 2004 (available at www.Paris21.org/knowledgebase).

OECD. 2007. *A guide to designing a national strategy for the development of statistics,* Paris 21 (available at www.paris21.org/pages/designing-nsds/presentation-events/index.asp?tab=doc).

Perrin, B. 2006. *From outputs to outcomes: practical advice and governments around the world*. Washington, DC, USA, World Bank and the IBM Centre for the Business of Government.

Ravallion. M. 2008a. *Evaluation in the practice of development*. Policy Research Working paper. Washington, DC, USA, World Bank.

Ravallion, M. 2008b. *Evaluating anti-poverty programs*. In *Handbook of Development* Economics, Vol. 4. Policy Research Working Paper Series. No 3625. Washington, DC, USA. World Bank.

United States Department of Agriculture (USDA)/National Agricultural Statistics Services (NASS). Crop Progress, (available at www.nass.usda.gov/Publications/National_Crop_Progress/terms_definitions.asp).

Verma, V., Marchant, T. & Scott, C. 1988. *Evaluation of crop-cut methods and farmer reports for estimating crop production. Results of a methodological study in five African countries*. London, UK, Longacre Agricultural Developments Centre Report.

World Bank. 2000. *Core welfare indicators questionnaire*. Washington, DC, USA (available at www4.worldbank.org/afr/stats/pdf/cwiq.pdf).

_____ 2003. *A user's guide to poverty and social impact analysis*. Washington, DC, USA.

_____ 2004. *Monitoring and evaluation, some tools and methods and approaches*. Washington, DC, USA, Operations Evaluation Department (OED).

_____ 2005a. *Agriculture and achieving the millennium development goals*. Washington, DC, USA, Agriculture and Rural Development Department.

_____ 2005b. *Gender issues in monitoring and evaluation in rural development: a tool kit*. Washington, DC, USA. June 2005.

_____ 2005c. *Results focus in country assistance strategies: the stocktaking of result- based CASs*. Washington, DC, USA, Operations Policy and Country Services.

_____2006a. *Conducting quality impact evaluation and the budget, time and data constraints*. Washington, DC, USA.

_____ 2006b. *IDA14 Results Measurement System (RS) update on commitments and implications*. Washington, DC, USA, Africa Management Team.

_____ 2006c. *Impact evaluation and the project cycle.* Washington, DC, USA, Thematic Group on Poverty Analysis Monitoring and Impact Evaluation.

_____2007. *How to build M&E systems to support better government.* Washington, DC, USA.

Wye City Group. 2000. *Handbook on rural households' livelihood and well-being. Statistics on rural development and agriculture household income.* Geneva, Switzerland (available at www.fao.org/statistics/rural).

Web sites:

FAO Database on food and agriculture
www.faostat.fao.org
FAO Fisheries resources monitoring system
www.firms.fao.org/firms
FAO Global information system on water and agriculture
www.fao.org/nr/water/aquastat/main/index.stm
FAO National food security statistics.
www.fao.org/faostat/foodsecurity/index_en.htm
FAO National Statistical Information System for Food and Agriculture
www.fao.org/statistics/countrystat
FAO Terminology
www.fao.org/faoterm
International Labour Organization
www.laborsta.ilo.org
Joint Monitoring Programme for Water Supply & Sanitation data WHO & UNICEF
www.wssinfo.org
Monitoring and Evaluation News
www.mande.co.uk
United Nations Classification of Functions of Government (COFOG) for agriculture.
www.unstats.un.org/unsd/cr/registry/regcs.asp?Cl=4&Lg=1&Co=04.2
United Nations Millennium Development Goals
www.un.org/millenniumgoals
United Nations Statistics Division, System of National Accounts 1993
www.unstats.un.org/unsd/sna1993/introduction.asp
United Nations Statistics Division, definitions.
www.unstats.un.org/unsd/cdb/cdb_list_dicts.asp
World Bank Evaluation, Monitoring and Quality Enhancement
www.worldbank.org/evaluation/
World Bank Monitoring & Evaluation Capacity Development
www.worldbank.org/oed/ecd/

World Bank PRSP Sourcebook
www.web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTPRS/0,,contentMDK:20177140~pagePK:148956~piPK:216618~theSitePK:384201,00.html
World Bank definition of proportion of the population below US$1/day poverty line
www.ddp-ext.worldbank.org/ext/GMIS/gdmis.do?siteId=1&contentId=Content_2&menuId=LNAV01HOME2

Annex 1: A List of core indicators

Annex 1 should be used in conjunction with Chapter 2 of the Sourcebook to help with the selection of appropriate indicators for monitoring ARD programmes. The list has been created through a participative process involving a number of different specialists and players. Initially, subject matter specialists were asked to use their expert knowledge to come up with the first basic list of indicators, paying particular attention to early outcomes indicators. Their suggestions were reviewed and merged to form the basic documentation for the five country reviews. In each country, a national workshop was organized in which national monitoring and evaluation (M&E) practitioners and statisticians were brought together to review the contents of the Sourcebook and to make recommendations based on their own practical experiences in the field. As a result of the workshops and feedback, the guidelines were extensively revised and the list of indicators updated.

The list contains 86 indicators, 19 of which are termed "priority indicators" (in green). They are the key indicators used for monitoring ARD programmes at the global level and to which all countries are encouraged to subscribe – not just for the purposes of feeding into the international monitoring systems, but also for serving as a standard core for monitoring the national ARD activities. The remaining indicators in the list constitute the extended list. They are not mandatory, but are intended to serve as a reference list that countries can use when deciding on what indicators they specifically wish to include in their own M&E systems, in addition to the priority indicators. The extended list therefore serves as a menu from which choices can be made. It provides suggestions and examples of indicators that can be used for monitoring and evaluation of a broad range of ARD activities.

The indicators are primarily outcome and impact indicators. Input and output indicators are not included. Each indicator includes the following items of information:

- sector/subsector, which shows the specific ARD activity for which the indicator is designed to be used
- class of indicator, which may be (i) early outcome; (ii) medium-term outcome; or (iii) long-term outcome, and indicate suitability for monitoring over different time periods;
- core data requirements, which are the data needed to construct the indicator;
- data sources – they may vary from country to country, but the list shows the most common source of information;

- technical notes to explain the critical concepts related to defining indicators.

Among the early outcome indicators are the service delivery indicators (access, use and satisfaction). The methodology for selecting and adapting these indicators to different subsector programmes is described in Chapter 2. The list in the Annex does not include all such indicators, but offers a few selective examples adapted to specific subsectors.

A large number of indicators come from survey data, which is disaggregatable. Although not specifically stated in the list of indicators itself, it should be standard practice that wherever possible, the indicators are disaggregated and shown by gender, by type of community (urban/rural) and by some measure of wealth ranking.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
A	**Sector-Wide Indicators for Agriculture and Rural Development**			
Early outcome				
1	**Public spending on agriculture as a percentage of GDP from the agriculture sector**	Government budget allocations, disbursements and spending on items related to agriculture; national agriculture value-added data	Ministry of Finance, National Accounts, National Planning Commission, country reports prepared by donors	This indicator demonstrates commitment to promoting agriculture. Before establishing a system to regularly compile the indicator, it would be necessary to clearly specify the related concepts, i.e. the items of public expenditure that would constitute the expenditure on agriculture. The definition for public spending on agriculture should follow the United Nations Classification of Functions of Government (COFOG) for agriculture. For further information, see: www.unstats.un.org/unsd/cr/registry/regcs.asp?Cl=4&Lg=1&Co=4.2.
2	**Public spending on agricultural input subsidies as a percentage of total public spending on agriculture**	Government budget allocations, disbursements and spending on items related to subsidies on agricultural inputs as well as total spending on agriculture	Ministry of Finance, National Planning Commission, country reports prepared by donors	Agricultural inputs should cover fertilizer, seeds, pesticides, water, etc. Direct cash subsidies paid to farmers and/or to input suppliers are more available and easier to estimate. Often, calculation of subsidies to agriculture is made with reference to border prices of inputs. Using border price criterion, the total subsidy paid by the government could be split into the portions going to farmers and non-farmers. The volatility of international prices makes such a comparison difficult over time. There may also be hidden subsidies due to differential prices for different sectors and cross-subsidization. Estimation of subsidies in big projects such as canal water supply schemes that supply water to agriculture at a nominal cost or no cost may require an in-depth study of the cost and price structure.
3	**Prevalence (percentage) of underweight children under five years of age in rural areas**	Anthropometric data	Anthropometric surveys	A high prevalence of underweight children in an area is an indicator of food insecurity in the area. Similar indicators based on “height for age” and “weight for height” are also commonly used.
4	Percentage of population who consider themselves better off now than 12 months ago	Data from specific household surveys	Special household surveys	This is an indicator of perception; hence “to be better” can mean different things to different people. This type of survey usually consists of completely open questions and does not refer only to “economic betterment”.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
Medium-term outcome				
5	**Food Production Index**	Area, production and yield data for major crops, livestock numbers and yields, felling rate, fishery production statistics	Statistical systems for crop, livestock, fishery and aquaculture. National Statistics Office and/or Minsitry of Agriculture	The category of food production includes commodities that are considered edible and that contain nutrients. Accordingly, coffee and tea are excluded along with inedible commodities because, although edible, they have practically no nutritive value. Separate indices for crop, livestock and fish production could be combined into a single Food Production Index on the basis of value of output. Before combining, the indices of value of output for each of the subsectors would need to be deflated to base year prices using suitable deflators. It will also be useful to disseminate the indices of production separately for crop, livestock and fish compiled on the basis of physical quantities. Yet another version of this indicator could be *"Per capita food production index",* which could be easily derived using population in the denominator. In countries where fibre (e.g. wool or jute) and biomass are important agricultural commodities, the index could be generalized to include food, fibre and biomass production. The commodities covered in the computation of indices of agricultural production by FAO are all crops and livestock products (www.faostat.fao.org). All products are covered, with the main exception of fodder crops. The fodder for the animal is a product of agriculture, but is also an input used within agriculture.
6	**Annual growth (percentage) in agricultural value added**	Values of inputs and outputs of agricultural commodities	National accounts wing of the National Statistics Office or CB	Once the time series on value added in agriculture sector is available, it may be preferable to use the trend growth rate, which is less affected by random factors operating in different years. Year-to-year growth rates are known to be affected more by base year figures. Calculations based upon triennium ending moving averages could also provide a good option. A comparison of annual growth and trend growth could provide a measure of impact of programmes. A variant of this indicator would be *"Annual growth (percentage) in agricultural value added per capita."* For calculating per capita, the base could be taken as the population dependent on agriculture. Another variant of this indicator is *"Value added in the agricultural sector per agricultural worker"*, which could be calculated taking agricultural labour force as denominator.
Long-term outcome				
7	**Rural poor as a proportion of the total poor population**	Household income and consumption estimates	Household income and consumption surveys	Countries should select the concept for poverty measurement (relative or absolute) that better suits their specific conditions.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
8	Percentage change in proportion of rural population below US$1 (Purchasing Power Parity) per day or below national poverty line	Income or consumption data (whenever possible, consumption is preferred to income for measuring poverty)	Household budget surveys and other surveys covering incomes and expenditure; World Bank Development Research Group	Proportion of population below US$1 per day is the percentage of the population living on less than $1.8 a day at 1993 international prices. The one dollar a day poverty line is compared to consumption or income per person, and it includes consumption from own production and income in kind. This poverty line has been fixed for measuring purchasing power across countries or areas and is often called an "absolute poverty line" or measure of extreme poverty. The international poverty line is based on the purchasing power parity (PPP) concept. The line is currently being revised using the 25 release of the International Comparison Programme on PPPs. For more information, see www.ddp-ext.worldbank.org/ext/GMIS/gdmis.do?siteId=1&contentId=Content_2&menuId=LNAV1HOME2.
9	Percentage of the population with access to safe/ improved drinking water	Total number of households; number of which: (i) are connected to water pipelines; (ii) have access to public water facilities; (ii) have protected wells, spring or rainwater	Administrative or infrastructure statistics from national and sub-national bodies, household surveys, Multiple Indicator Cluster Survey, Demographic and Health Surveys, Living Standard Measurement Surveys, Joint Monitoring Programme for Water Supply & Sanitation data from WHO and UNICEF. See www.wssinfo.org.	A population will have access to safe drinking water if a safe and dependable source is located within a reasonable distance. The reasonable distance should be quantified as per local conditions. Sources of safe drinking water will include piped water, public tap, borehole or pump, protected well, protected spring or rainwater. Improved water sources do not include vendor-provided water, bottled water, tanker trucks or unprotected wells and springs.
10	Consumer Price Index for food items	Consumer price index disaggregated by categories	National Statistics Office	Consumer price indices are grouped by type of items and "food" is usually one of them. The current practices of each country concerning what must be considered inside this category should be preserved. Nevertheless, some broad indication must be made to allow international comparisons; for example, "the food group must contain food and beverage goods".
11	Agricultural exports as a percentage of total value added in agriculture sector	Trade statistics, national accounts statistics	National Trade data; United Nations Statistics Division Key Development Data & Statistics, FAO, FAOSTAT database (www.faostat.fao.org)	To allow international comparisons, the definition of "Agricultural goods" in FAOSTAT metadata should be adopted.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
12	Proportion of under-nourished population	Data from household budget surveys	FAO National Food Security Statistics (www.fao.org/faostat/foodsecurity/index_en.htm)	Undernourishment is defined as a condition that results from consuming too little food over a period of time. Like malnutrition – the result of a diet that is lacking in certain nutrients (such as protein or vitamins) – undernourishment is common in poor countries. The number of undernourished people is estimated on the basis of national norms of daily calorie consumption using data from consumption surveys. Norms in the country may vary according to rural and urban areas or gender.
13	Producer Price Index for food items	Prices received by domestic producers of food items	Prices surveys; www.faostat.fao.org	Wholesale prices of agricultural commodities are often available. There may be a need to adjust these prices to obtain the producer's price. To build an index of these prices, due consideration should be given to the weight to be assigned to different commodities and the base year. The scope of such an index could be all agricultural commodities including food and non-food crop and livestock products. The National Statistics Office should be involved in the process of building such an index.
14	Ratio (or proportion) of arable land area to total land area of the country	Land use data	National Agricultural Statistics available from the Ministry of Agriculture and National Statistical Office; FAO agricultural resources statistics (www.faostat.fao.org)	Special attention must be given to the definition of arable land. The category as defined in the FAO WCA 21 could be adopted. See para. 11.22 of FAO Statistical Development Series No. 11.
15	Change (percentage) in unit cost of transportation of agricultural products	Cost and volumes of agricultural products transported by mode of transport and region	Unions of transport providers; national transport companies	This cost will usually be calculated per km, per tonne. It may vary significantly depending upon mode of transport, region of transport and the distance travelled.
16	Percentage of rural labour force employed in agriculture	Total number of days/persons (rural household members) worked in agriculture, by type of work (paid/unpaid; permanent/occasional) during a specified period.	Labour force surveys, Population Census, Agricultural Census, International Labour Organization (www.laborsta.ilo.org)	An important component of the labour force for agriculture includes family workers who are mostly unpaid and often not counted. One way to measure the rural labour force is through a household survey that specifically collects information on the number of days or average hours each household member works on its own holding without explicit wages and the number of days or hours they work for wages, separately for farm and non-farm activities. Many will do both, i.e. work on their own holding and somewhere else for wages. It is also important to know the demographics of the rural labour force; to answer questions like "how a programme to improve school attendance will affect the rural labour force. This would need data on labour force by age.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
17	Percentage of rural labour force employed in non-farm activities	Activity status of rural household members, time worked in each activity	Farm labour data from national Census of Agriculture	See Technical Notes on Indicator at Serial No.16.
18	Percentage of the labour force underemployed or unemployed	Economically active population, employment and unemployment data on hours/ days of work	National Labour Statistics, labour force surveys, International Labour Organization (www.laborsta.ilo.org)	See Technical Notes on Indicator 16. Countries usually follow their national definition for measuring "underemployment" and "unemployment". For international comparison, these definitions would need to follow concepts proposed by the International Labour Organization. The definitions related to employment usually follow an approach based on "Current/daily Status" , "Weekly Status" or "Usual Status". See also www.laborsta.ilo.org for details.
19	Annual growth rate of household income in rural areas from agricultural activity (percentage)	Household income by source and groups of activities	Annual income-expenditure surveys	In the absence of this indicator, a comparison of wage rate for agricultural workers and non-agricultural workers in rural areas could provide a proxy (not a substitute) indicator. In some situations, the incomes from the two subsectors may show a completely different trend since incomes also depend on the duration of employment in addition to wage rate.
20	Annual growth rate (percentage) of household income in rural areas from non-agricultural activity	Household income by source and groups of activities	Annual income-expenditure surveys	See Technical Notes on indicator at Serial No. 19.
B	**Specific Indicators for Subsectors of Agriculture and Rural Development**			
1.	**Crops (inputs and services related to annual and perennial crop production)**			
Early outcome				
21	Indicators of access, use and satisfaction with respect to services related to sustainable crop production practices, technologies and inputs, for example: (i) percentage of small-scale farmers who know about sustainable crop production practices (ii) percentage of farmers who applied/ purchased the recommended package of inputs last season (iii) percentage of farmers who adopted sustainable crop practices in their farms	Total number of farmers and area of their holdings; number of farmers who know/use sustainable crop production practices; total area under sustainable crop practices	Agricultural extension services; sustainable crop production certifying bodies; agricultural/ environmental services; analysis and studies made on the basis of data available from agricultural census and surveys	From among a number of local crop production practices, certain practices will be classified by the national agricultural research and extension system as being "sustainable". Usually, there would be an attempt by the extension system to promote such practices among farmers. The special survey to collect data for this indicator will be directed towards use of such practices. In addition to asking a direct question on sustainability in a special survey, one could also derive these types of indicators using information on input use, crop rotation practices, etc. available in a regular agricultural survey.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
Medium-term outcome				
22	**Change (percentage) in yields of major crops of the country**	A time series of crop yields per unit of land area for major crops or crops covered by a specific programme	Objective crop yield measurement surveys or other forms of assessment	Because of high year-to-year fluctuations, particularly in rainfed areas, statistically significant trends may not be discernible until an 8-9 year series is established. In agricultural statistics, the concept of yield has been generally used to represent the average amount of produce obtained per unit of crop area. In some cases of tree crops, the concept of yield covered the average amount of produce per tree. For further information, see FAO Economic and Social Development Paper No. 22, para. 61-63: "Estimation of crop areas and yields in agricultural statistics".
Long-term outcome				
23	Yield gap between farmers' yields and on-station yields for major crops of the country	Yield expected to be achieved at farmers' field with recommended package of inputs and actual yield as estimated through the agricultural statistics system	Crop surveys and information on potential yield of a crop variety indicated by agricultural research and extension system	Under normal weather conditions, yield-gap is an indicator of available potential for growth. The potential may remain unrealized due to various factors, such as unaffordability/ unavailability of inputs and/or the difference between controlled conditions in a research institution and uncontrolled environment on the farmers' fields.
24	Percentage of total land area under permanent crops	Land use data	Agricultural census and current agricultural surveys	One could also extend this indicator to monitor composition of total arable agricultural land, which includes land under permanent crops, temporary crops and permanent meadows, as well as growth of any of the individual constituents of total arable land.
2. Livestock				
Early outcome				
25	Indicators of access, use, satisfaction with respect to livestock services, for example: • Percentage of livestock owners in contact with livestock officer in the last month • Percentage of livestock owners using veterinary services within the last month • Percentage of livestock owners satisfied with the quality of livestock services	Total number of livestock owners; number of which were visited by a livestock officer during the last month; number of which were satisfied with the quality of services received by type of service.	Veterinary extension services; analysis and studies made on the basis of data available from livestock census and current surveys; specific livestock surveys	A number of veterinary services, e.g. dipping, vaccination and artificial insemination are provided by government agencies responsible for livestock development. These services could also be purchased from the private sector. In designing the survey, care needs to be taken to specify to which particular service or sector the survey refers – public, private or both.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
Medium-term outcome				
26	**Annual growth (percentage) in value added in the livestock sector**	Values of input used in livestock sector and output produced	National Accounts wing of the National Statistics Office	Standard concepts used for national income accounting should be followed. Gross value added is the value of all goods and services produced during a production period but not immediately used up in the production process of that period. Hence, gross value added represents the value of all goods and services available for the different uses other than intermediate consumption. (Intermediate consumption consists of the goods and services consumed in the production process, excluding the consumption of fixed assets.) Gross value added = output – intermediate consumption. The formula for estimating the output of livestock in general is based on the following relationship: Output of live animals + Imports= Animal slaughtered or died of natural causes + exports + change in animal stock. For more details, see System of National Accounts (1993, para. 6.94-6.1): www.unstats.un.org/unsd/sna1993/introduction.asp and *A System of Economic Accounts for Food and Agriculture* (FAO, 1996; www.fao.org/docrep/W1E/W1E.htm).
Long-term outcome				
27	Livestock birth rate	Number of births during the last year and number of female animals of reproductive age for each species	Periodic livestock surveys and estimates prepared by livestock specialists	To be compiled separately for each species.
28	Percentage increase in yield per livestock unit	Per animal yield of milk, eggs, meat, wool etc., separately for each species	Livestock surveys; FAO Yield Livestock data (www.faostat.fao.org)	Yield refers to milk, eggs, meat or wool per livestock. As there may be more than one product from same animal, this indicator needs to be compiled separately for each species and product. However, the yield of the product that is the main purpose of production is the most relevant yield. Seasonality in yield is important in some products. Therefore, for comparison of growth, the use of comparable periods, which would vary from country to country, must be stressed.
29	Percentage change in livestock values	Livestock number and unit prices of livestock by species and breed	Department of Livestock, National Accounts wing of the National Statistics Office	The value of stocks of different type of livestock is considered capital stock or intermediate stocks, depending on the purpose of rearing. Such data is usually estimated during the preparation of National Accounts.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
3. Fisheries and Aquaculture				
Early outcome				
30	Indicators of access, use, satisfaction with respect to fisheries/aquaculture services, for example: • Percentage of fishing communities in contact with a fisheries officer in the last month • Percentage of rural communities that constructed a fish pond in the last year • Percentage of fishers satisfied with the quality of fisheries services	Total number of fishing communities/fishing households; the number of which received a visit from a fisheries officer during the last month, the number of which constructed fish ponds in the last year; and the number of which were satisfied with the quality of services received by type of service.	Stakeholders surveys; Extension wing of the Department of Fisheries and Aquaculture	In some countries, small-scale fish ponds are advocated for households instead of community aquaculture pond. In such cases, "*Number of households having established a private fish pond*" will be a substitute indicator.
31	Water use per unit of aquaculture production	Production from aquaculture, water used in aquaculture production, number of aquaculture units and their water holding capacity	Special surveys of aquaculture production units; Department of Fisheries	This indicator will not be relevant for aquaculture activities carried out in running water or the sea. The indicator purports to measure resource use efficiency of water used for aquaculture. It could be calculated both in terms of physical quantities and values, particularly if more than one product is being produced by the same production unit or when the data from companies engaged in production of different products are to be merged.
Long-term outcome				
32	**Capture fish production as a percentage of fish stock (or a rating of the state of major capture fish stocks relevant to exports and local food)**	Scientific estimates of fish stocks and exploitation rates or perceptions/assessment of community of fishers on increasing or decreasing fish stock	Institutions involved in estimation of fish resources; stakeholder survey on perceived state of fish stocks	This indicator particularly refers to capture fisheries in natural water bodies and not the culture fisheries that usually follow a well-established culture and catch cycle. For planning an assessment survey, one may use seven levels for the fish stocks state ratings, viz., not known, underexploited, moderately exploited, fully exploited, overexploited, depleted and recovering; For details, see ftp://ftp.fao.org/docrep/fao/7/y5852e/y5852e.pdf. A bi-dimensional rating of state of stocks based on (i) exploitation rate and (ii) stock abundance could also be used. For details, see www.firms.fao.org/firms.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
33	Share of small-scale fishers in the production of fish	Average fish price, number of small-scale fishers or aquaculture producers, days of fishing, average weight per day of fishing	National fisheries surveys; estimates prepared by National Accounts wing of the National Statistics Office	This indicator could be calculated for both inland and marine fishery. If many different types of fish with varying prices are being produced in the country, it may be appropriate to compile this indicator on the basis of total value of the production. There may not be a universally accepted criteria for defining "small-scale fishing", which is usually done on the basis of type of boat and fishing gear, as well as socio-economic factors such as ownership of boat, crew size and marketing methods. Often, the term "artisanal fishery" is used to describe small-scale fishing.
34	Fishing quota (percentage of total permitted catch) earmarked for local fishing communities as rights	Quantity of catch permitted, by species, season and fishing community, according to local practices and regulations	Fishing regulatory bodies	This may be applicable to both coastal fishing and fishing in inland ponds and lakes controlled by the government.
35	Annual change (percentage) in production from aquaculture farms	Quantity and average unit price of different products from aquaculture	Department of Aquaculture and the National Statistics Office	Aquaculture products include weeds, plants, animals and even ornamental and medicinal products. Efforts should be made to cover all such products being cultured in-land and in the sea. It will be useful to compile this indicator in terms of both quantities and values. The value of a diverse group of products could be pooled to form a single indicator. It would be desirable to compile this indicator at the country, region, and district level, or even at lower levels.
4. Forestry (developing, caring for or cultivating forests; management of timber production)				
Early outcome				
36	Indicators of access, use, satisfaction with respect to the forestry services: • percentage of communities aware of the activities of forestry services in their area • percentage of communities involved in sustainable forest management • percentage of communities planning to expand area under sustainable forest management	Total number of communities; the number of which are aware of forestry services in their area; the number of which are involved in sustainable forest management; area under sustainable forest management and plans for the next year per rural community	Stakeholders surveys	Usually, agricultural surveys and censuses do not have enough information on sustainable forest management; therefore ad hoc surveys are needed.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
37	Employment in forestry-related activities (full-time equivalents)	Data on paid employment and self-employment	Ministry / Agency responsible for forestry, the National Statistics Office, special surveys	Employment should ideally include both paid employment and self-employment. However, data on the latter is often unavailable in national statistics. For paid employment, data from national statistics (ISIC/NACE Rev. 4 Activity A2) can be used. For definitions, see www.unstats.un.org/unsd/cdb/cdb_dict_xrxx.asp?def_code=388.
38	Value of removals of wood and non-wood forest products (selected currency)	Authorizations and licenses granted, statistics on removals, trade statistics	Ministry / Agency responsible for forestry, special surveys	The value of the removals of a product indicates its socio-economic importance, as long as they are harvested within the framework of sustainable management.
39	Value of services from forests (selected currency)	Value of carbon sequestration, tourism, water supply, etc.	Ministry / Agency responsible for forestry, Ministry of the Environment, special surveys	Forests provide a number of environmental services, many of which are difficult to assess in terms of value. However, these services can be of great importance, particularly for individual projects, and in such cases, efforts should be made to assess them.
Medium-term outcome				
40	Area of forest under sustainable forest management (hectares)	Area with forest certification, area with forest management plan, local (documented) knowledge, time series if available	Ministry of Environment, Ministry/ Agency responsible for forestry, certification bodies	Although there is not yet any internationally agreed definition of "sustainable forest management", this information is required by many international agencies like CBD, UNFF, ITTO and FAO. When this indicator is compiled, the criteria used for its estimation should be documented.
Long-term outcome				
41	**Proportion (percentage) of land area covered by forest**	Area of forest, land area	Ministry/Agency responsible for forestry, geographical institute	Forest is defined as "land spanning more than .5 hectares with trees higher than 5 metres and a canopy cover of more than 1 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use" (FAO).
42	Annual growth (or percentage change) in rural household income from forest-related activities	Composition of rural household incomes	Household income surveys for rural areas	Only such forest-related activities that are a source of livelihood for the rural population should be considered. In the absence of this indicator, a comparison of wage rate for workers in forest-related activities and other workers in rural areas could provide a proxy (not a substitute) indicator.
43	Growing stock per hectare (m^3/ha) of forest	Area of forest, growing stock (volume)	Ministry/Agency responsible for forestry	Growing stock is the volume of the standing trees and it can be converted to biomass and carbon stocks using conversion factors provided by IPCC. The growing stock per hectare indicates if the forests in a given area have lost or accumulated stocks.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
44	Rate of deforestation (percentage)	Information on area under forest cover, area reforested (additions to forest stock) and deforestation (deletion from forest stock) during the year	Ministry of Environment and Forests	Deforestation is the conversion of forest to another land use or the long-term reduction of the tree canopy cover below the minimum 1 percent threshold. It implies the long-term or permanent loss of forest cover and the transformation into another land use. It includes areas of forest converted to agriculture, pastures, water reservoirs and urban areas. It also includes areas where the impact of disturbance, over-utilization or changing environmental conditions affects the forest to an extent that it cannot sustain a tree cover above 1 percent threshold. However, it excludes areas where the trees have been removed as a result of harvesting or logging, and the forest is expected to regenerate naturally. See www.fao.org/faoterm. Another related indicator would be *"Ratio of land area forested during the last year to total land area"*.
5. Rural Micro and SME Finance				
Early outcome				
45	Indicators of access, use, satisfaction with respect to rural finance services, for example: • percentage of the rural population eligible to obtain a business loan • percentage of users who are satisfied with the banking services	Total number of rural households; number of which are eligible to benefit from rural finance services by type of loan requested; and number of which are satisfied with the quality of banking services.	Central Bank or Lead Commercial Banks active in an area; special survey	In addition to eligibility criterion, access to banking services could be measured in terms of the extent of the network of banking institutions or distance at which a banking branch is located. While the statistics on use of banking services would be available from a banking institution, a survey of users may be desirable for assessing satisfaction of services. For planning such a survey, a list of customers available with banking institutions could serve as a sampling frame. See also Indicator of use at Serial No. 46.
46	**Percentage of the rural population using financial services of formal banking institutions**	Total number of rural households; number of which requested/accessed credit and/or savings from formal banking institutions	Central Bank or Lead Commercial Banks active in an area, population census, special survey	Use of financial services refers to loans, credit cards and deposit accounts of different types, etc. The agricultural census and surveys are often a source of information for this indicator with reference to households of agricultural holders. Agricultural censuses usually have a section on agricultural credit where access and use by type of credit institution are reported.
47	Percentage of bank branches that are located in rural areas	Spatial distribution of bank branches	Central Bank or Lead Commercial Banks active in an area	GIS and agricultural atlases can be useful in collecting and representing such data.
Long-term outcome				
48	Percentage of total savings that are mobilized from rural areas	Mobilization of savings from different areas	Central Bank or Lead Commercial Banks active in an area	Standard banking concepts should be followed.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
49	Percentage of rural population using non-bank financial services	Total number of rural households; number of which requested/accessed non-banking financial services	Insurance and leasing companies and special surveys	Non-banking financial services refer to leasing and insurance.
50	Recovery rate of rural credit	Volume of loans by type of credit and rate of recovery in a specific period	Central Bank or Lead Commercial Banks or refinance institutions active in the area	Rate of repayment is a common indicator used by banking institutions for management of their operations. Usually, the recovery rate may be different for different institutions as well as for different category of loans. However, a range of recovery rates by types of credit is a useful indicator for monitoring purposes.
6. Agricultural Research and Extension				
Early outcome				
51	Indicators of access, use, satisfaction with research and extension advice, e.g. • percentage of farmers having knowledge of a specific technology being disseminated by extension system • percentage of farmers trying the specific technology advice from extension systems on their farms • percentage of farmers who were satisfied with the specific technological recommendations of the extension system and judged it beneficial, with or without adaptation	Total number of farmers; number of which had knowledge of a specific technology disseminated by the extension services; number of which applied the technologies suggested by extension services; and number of which were satisfied with the quality of the extension services	Special surveys	Although research and extension services are well connected, a distinction must be made between them in order to measure the quality of service. It may also be important to make a distinction between different channels of extension information as well as different service providers, e.g. public and private. Visits of extension workers to farmers lead to two-way communication. The examples given here actually refer to technological advice rather than the service provided by extension workers in bringing this advice to the farmers. A survey of the farmers to assess the services offered by the extension system would include indicators such as: (1) percentage of target farmers visited by extension workers in last two weeks; and (2) percentage of farmers satisfied with extension services in general. It should be noted that during the visits, not only is the new technology extended to the farmers, but also, feedback on the technology is obtained. Since indicator (1) is to measure the "degree of activity" of the extension system, the number of contacts of extension workers with the farmers should be taken into account, without considering the purpose of the visit.
52	**Public investment in agricultural research as a percentage of GDP from the agriculture sector**	Budget allocations to agriculture research institutions, GDP for agriculture	Ministry of Finance; National Accounts Statistics	Since the same agency is very often involved in both research and extension, public expenditure on these two activities is often reported together. Care needs to be taken to separate the two types of expenditures. In the absence of any other method, one could apportion the expenditure in the ratio of number of research and extension personnel.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
Long-term outcome				
53	Percentage change in yields resulting from improved practices, for major crops of the country	Crop yield data for major crops	Current agricultural statistics or assessments based on interviews of farmers	It is often difficult to isolate the effect of several factors contributing to yield, e.g. variety of seed, fertilizer, water and weather. In order to establish if the observed increase in yield was caused by the improved practices, some type of experimental design for counterfactual comparison must be set. This indicator would refer to a specific project where the beneficiary farmers could be directly asked about their assessment of gain in yield.
54	Change in farmer income as a result of new technologies (by gender)	Yield before and after introduction of new technology, prices of outputs, distribution of agricultural land, area dedicated to new technology	Special studies on improved agricultural practices	Wherever detailed income surveys are not being conducted, this indicator may be compiled on the basis of the assessment and opinions of a target group of farmers.
7. Irrigation and Drainage (services related to water use in agriculture)				
Early outcome				
55	Indicators of access, use, satisfaction with respect to irrigation and drainage services, for example: • percentage change in the proportion of farmers with access to a functioning (reliable and adequate) irrigation and drainage network • percentage change in the number of users	Total number of farmers; number of which have access to a reliable/adequate irrigation and drainage network; number of which use the irrigation and drainage network	Agricultural census; other crop related surveys or water user survey	See Technical Notes on indicators at Serial No.56 and 57. The actual number of beneficiaries from a canal depends upon the extent to which it is a reliable source of water. The adequacy of irrigation should be measured in the cropping season when irrigation is most needed.
56	**Irrigated land as percentage of crop land**	Total crop land area; irrigated crop land area	Agricultural census and current agricultural surveys	Irrigation refers to purposely providing land with water other than rain for improving pasture or crop production. Irrigation usually implies existence of infrastructure and equipment such as canal, pumps, sprinklers or a localized watering system. It also includes manual watering of plants using buckets, water cans or other devices. Uncontrolled land flooding by overflowing or rivers or streams is not considered irrigation. For further information, see FAO Statistical Development Series No. 11 (para. 11.68-11.9): *A system of integrated agricultural censuses and surveys,* Vol. 1. Countries should formulate a precise definition of irrigated and drained land by referring to FAO statistical guidelines. Land area that receives at least one controlled irrigation a year may be considered irrigated. The concept of "command area for irrigation canals" often leads to an over-estimate of the number of beneficiaries

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
57	Percentage of users who report a significant increase in crop yields as a result of the provision of irrigation and drainage services	Crop yield data for irrigated crops; crop yield data for the same crops cultivated in similar conditions without irrigation	Agricultural census; other crop-related surveys or water user survey	See Technical Notes on indicators at Serial No. 55, 56 and 62.
58	Service fees collected as a percentage to total cost of sustainable Water User Association (WUA) activities and functions	Total WUA budget; part of which came from collected fees.	Special studies on financial aspects of WUA	Usually, a considerable proportion of WUA's income should come from service fees, the information on which should be available in administrative records
Long-term outcome				
59	Percentage change in average downstream water flows during dry season	Monthly average downstream water flows	Records of project authorities	Regional records are usually also available.
60	Percentage change in agricultural value added created by irrigated agriculture	Area benefiting from a project, area under major crops of the command area of the project, yield of crops before and after availability of irrigation, prices of the crops sown in the command area	Special studies	The study may need to confine to command area of an irrigation project.
61	Percentage of irrigation schemes that are financially self-sufficient	List of irrigation schemes, their running costs and revenue collected	Ministry of Water Resources	Financial self-sufficiency here refers only to operation and maintenance costs.
62	Percentage increase in cropping intensity	Area equipped for irrigation data, crop yields in irrigated area	Census of Agriculture; current agricultural surveys and related surveys	This indicator is aimed at judging the effects of irrigation projects. There are other possible proxy indicators for monitoring effects of an irrigation project, viz., yield of crops sown in the area or changes in choice of crops itself. See, for instance, Indicators at Serial No. 23 and 57. It should be noted, however, that the cropping intensity is highly positively correlated with the extension of irrigation facility, whereas gains in yield and choice of crops could be attributed to other factors as well.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
8. Agri-business (agricultural marketing, trade and agro-industry)				
Early outcome				
63	Indicators of access, use and satisfaction with respect to agribusiness and market services, e.g. • percentage of farmers aware of market price and information services • percentage of farmers using market price and information services • percentage of farmers who are satisfied with agri-business and market services	Total number of farmers; number of which are aware of market prices and information services; number of which used market price information and/or market services; and number of which were satisfied with market information services received.	Stakeholders surveys	Surveys could be complemented by studies on the means of communication for each region, their audience, etc.
64	Percentage change in number and value of activities managed by agro-enterprises	Number and value of transactions managed by agro-enterprises, by type of business	Enterprise survey	In order to get the information from an enterprise survey on a regular basis, it may be necessary to review the definition of enterprise so as to include small-scale rural enterprises. This indicator could be compiled on the basis of information that is usually maintained as a sampling frame for the enterprise survey.
65	Proportion (percentage) of agro-enterprises adopting improved/ certified hygiene/food management system	Number of agro-enterprises by type of business. Number of agro-enterprises certified as following specified phytosanitory measures and food quality standards	Ministry of Industry; hygiene and food certification bodies	A precise description of "improved/certified hygiene/food management system" is needed. This may be done following national standards for food safety and processing.
Medium-term outcome				
66	**Change (percentage) in sales/ turnovers of agro-enterprises**	Benchmarks and subsequent data on enterprises covering sales, cost structures, and gross and net profits	Direct data collection through special surveys, including enterprises and a control group of enterprises	Standard accounting concepts should be followed.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
Long-term outcome				
67	Percentage change in number of agricultural inputs outlets	Number of retail dealers for each type of inputs	Ministry of Agriculture; companies marketing agricultural inputs	Such information is usually maintained by the marketing division of input supplier companies. However, a system of reporting may need to be established to compile this information on a periodic basis.
68	Percentage increase in private sector investments in agriculture	Private capital formation in agriculture and non-agriculture sector in rural areas	National accounts statistics; special surveys to assess private capital formation in an area	The investment in rural areas is done by both the public sector and the private sector. Public sector investment stimulates private sector investment. Estimates of both public and private capital formation are compiled by the organization responsible for National Accounts. Often, however, it may be difficult to find these estimates at the subnational level or disaggregated by rural and urban areas, whereas project monitoring tends to focus on a specific area. In such a situation, it may be necessary to carry out a special survey, e.g. Rural Investment Climate Survey (RICS) focusing on households and enterprises in the specified area. See an example of RICS in Indonesia at www//web.worldbank.org. Such indicators could also be generated using multipurpose surveys such as Living Standard Measurement Studies (LSMS).
69	Percentage increase in market share of cooperatives/agribusiness enterprises	Time series of value of transactions of cooperatives/ agribusiness enterprises, by product or group of products.	Ministry of Industry, special studies	This may be compiled for specific product or for a group of products. The activities to be covered in the agri-business sector would need to be precisely defined.
C Indicators for Thematic Areas Related to Agriculture and Rural Development				
1. Community-based Rural Development				
Early outcome				
70	Indicators of access, use, satisfaction with respect to services provided by community-based rural development organizations, for example: • percentage of members of community/producer associations reporting increased production or profits as a result of membership	Total number of farmers who are members of community-based rural development organizations; number of which reported increased production or profits as a result of membership, and number of which were satisfied with the with services received.	Special surveys of households and/or members of the organization	The increase in production/profits directly ascribed to membership is difficult to measure. The indicator should be estimated from subjective perception of surveyed farmers. The membership records of community organizations and demographic information would provide a sampling frame for the survey,

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
71	Percentage of farmers who are members of community/producer organizations	Total number of farmers; number of farmers who are members of community/ producer organizations	Special surveys to directly ask households or indirectly compiled on the basis of the membership record of community/producer organizations and demographic information	Community/producer organizations such as cooperatives or self-help groups may be working in one or more of the activities related to production and marketing. Their activities may include: input supply, water use, credit supply and marketing of produce, among others. Indicators on groups/associations/ organizations for specific activities will also result in useful indicators for monitoring community development as will the development of the specific ARD subsector. For example, information on "self-help groups working in micro-credit" will be an indicator for the rural credit subsector.
72	Proportion of community/producer organizations capable of meeting the production and marketing needs of their members	Total number of community/ producer organizations; number of which were rated by their members as meeting their needs, by type of membership.	Stakeholders surveys and assessments	In order to qualify the organization as meeting the production and marketing needs of its stakeholders, it would be necessary to establish some quantitative lower bound on the percentage of stakeholders who confirm this capacity. For example, if more than 8 percent of surveyed stakeholders admit that "Organization A" is capable of meeting their production/marketing needs, then "Organization A" qualifies.
73	Proportion of community organizations (POs)/NGOs with functional internal system of checks and balances	Total number of community organizations; number of which meet the pre-established standards	National Registry of NGOs; governing laws, procedures and accounts of NGOs	This would need special studies that will make a subjective assessment of each of the NGOs based on their accounts, rules and laws of management.
74	Percentage change in number of community associations exercising voting power in local government budget	Time series of total number of community associations; number of which exercised voting power in local budget allocation	Survey of community associations	It should be useful to stratify the community organizations according to size of their budget (large, medium, small).
Long-term outcome				
75	Percentage increase in number of local enterprises in rural area	Series of total number of local enterprises, by region	Enterprise survey, special survey	A local business would be defined that is mainly owned by residents of the area. These businesses would usually be family-owned enterprises operating in rural areas and agro-industries. The framework would need to be more precisely defined at the time of planning the survey. The scale of operation of the enterprise to be covered by the survey should be fixed at the national level.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
2. Natural Resource Management				
Medium-term outcome				
76	**Withdrawal of water for agricultural as a percentage of total freshwater withdrawal**	Area equipped for irrigation; area under different crops under irrigated and rainfed conditions; irrigation intensity and water requirement ratios of different crops; number of irrigations actually provided by farmers in season (on an average in an area); estimates for per capita water consumption by humans and animal; information on lift water irrigation devices (e.g. wells), etc.	National Ministry of Water Resources; special studies using crop and irrigation data from agricultural census/ surveys to estimate use of water in agriculture, per capita consumption by humans and consumption by industries	In order to ensure comparability over time, it would be necessary to clearly specify the related concepts prior to establishing a system to regularly compile data for the indicator. This indicator should take into account the use of both surface and groundwater. It may be necessary to establish the methodology through a working group of local experts and consulting internationally established methods. See also: www.fao.org/nr/water/aquastat/water_use/index.stm; and www.fao.org/nr/water/aquastat/data/query/index.html.
77	**Proportion (percentage) of land area formally established as protected area**	Information on land area of formally protected areas; total land area of the country or project area	Ministry of Environment	IUCN defines six categories of protected areas, ranging from strict nature reserves to managed resource protection areas. Countries usually assign their protected areas to one of these categories. Depending on the objectives of an evaluation, all or a subset of these categories can be used.
78	**Change (percentage) in soil loss from watersheds**	Area that has become uncultivable or has witnessed substantially reduced yields due to soil erosion and total area of watershed	Watershed authorities	Soil erosion is one type of soil degradation. According to FAO WCA 21, "Soil erosion is the displacement of soil material by running water, rainfall, wind or other factors resulting in decline of arable layers" (see para. 11.65). This indicator should be compiled on the basis of a quantitative measure of area affected by soil loss.
Long-term outcome				
79	Change (percentage) of farm l and under risk of flood/drought	Time series of information on drought/flood-prone areas of farm land and total farm land area	Ministry of Agriculture, Ministry of Environment	The basis of this indicator will be national assessments of drought/flood-prone areas.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
3. Land Policy and Administration				
Early outcome				
80	Percentage of land area inventoried	Total land area and total registered land area	Census of Agriculture, Land/Cadastral Register	It is important to consider the date of the cadastre. In some countries where cadastral registers are outdated, the indicator should be accompanied by references to the date of cadastral register to avoid confusion.
81	**Percentage of land area for which there is a legally recognized form of land tenure**	Total land area and area over which holders have a well-defined right to access and use	Agricultural census; Land Registration Authorities	There may be many types of land tenure forms recognized by a local court of law for adjudication. Such land tenure forms would be considered "legal". What is "legal" in one country may be "non-legal or illegal" in another. The intention here is to assess the area of land over which owners or tenants have a stable right to access and use. For more details, see concepts recommended in the FAO World Programme for Agricultural Census 21 (pp. 8-81) at: www.fao.org/es/ess/census/wca21.asp. Note that the information obtained from agricultural census would refer only to the land operated by the agricultural holders. Such land which is not being operated for agriculture will not be covered in the data obtained from agricultural census.
Long-term outcome				
82	Share (percentage) of land over which there are disputes	Total area under dispute on a precise reference date; total land area at that date	Land/Cadastral Register; Land Dispute Settlement Authorities/ Courts	This is an indicator of the state of functioning of the land regulation laws and the related administrative system. In the absence of area under disputes, the number of disputes could be a substitute indicator.
83	Percentage of agricultural households that have legally recognized rights to land	Information on land tenure	Population Census and Agricultural Census, special survey in project area	An agricultural household is one whose primary source of income is agriculture. An agricultural operational holder is one who manages some land or livestock as his/her own enterprise. Most agricultural holdings need land to carry out their operations, except perhaps for nomadic livestock holdings. The indicator thus gives the percentage of such household whose main source of income is agriculture and who operate the land with a legal right (legal or customary ownership or legal tenancy). Another complementary indicator would be *"% change in population of landless agricultural household"* or *"landless agricultural labour"*. For calculating these indicators, one would need to fix the cut-offs on size of land below which the household would be considered landless. Such a cut-off may be decided on the basis of data available from agricultural census.
84	Percentage change in number of formal land transactions (quarterly or yearly basis)	Series of quarterly/yearly formal land transactions; estimate of total number of land transactions.	Land registration authority; land/ cadastral register	Since it is difficult to assess the total number of non-formal land transactions, the indicator focuses only on formal transactions.

Sl. No.	Indicators	Core Data Requirements	Data Sources	Technical Notes
85	Change (percentage) in land access for women and minority groups	Percentage of land owned or under owner-like possession by different groups including women and minority groups	Agricultural Census, land/cadastral register	Usually, the agricultural census has information about the operational holder and not the owner of the land.
4. Policies and Institutions				
Long-term outcome				
86	Ratio of average income of the richest quintile to the poorest quintile (%) in rural areas	Average income per quintile in rural areas	Household budget or income surveys	A comparison of this indicator for rural and urban areas, at least at the national level, as well as by gender could provide meaningful insight into income inequalities.

Annex 2: Country Case Studies

Summary of country studies and of ARD indicators currently in use in each country

The Annex is divided into two parts. Part 1 consists of a summary of the five country case studies that were used as an integral part of the validation process. Part 2 consists of tables showing the indicators currently in use in each of the five countries.

Part 1 – Country studies

Country study 1 – Cambodia

The M&E policy environment – There is presently a favourable environment for putting in place a functional monitoring and evaluation (M&E) system in Cambodia. The current National Strategic Development Plan (NSDP) provides clear policy guidelines for the integration and use of an M&E system as a tool for systematically tracking progress of strategic programmes and actions towards achieving goals and objectives of the plan.

Institutional supports for M&E – The Ministry of Planning (MoP) was designated as the lead ministry responsible for: preparing the overall framework outlining the methodology; determining the frequency of reporting; coordinating activities; and consolidating and preparing the NSDP Annual Progress Report. The line ministries/agencies are responsible for monitoring and collecting input and output indicators, while the MoP is in charge of monitoring and evaluating outcome indicators through its periodic surveys undertaken by the National Institute of Statistics (NIS), the only legally and technically competent agency for the collection, processing, management, and presentation of various data on the country. In general, almost all government line ministries/agencies, including the Ministry of Agriculture, Forestry and Fisheries (MAFF) and the Ministry of Rural Development (MRD), have M&E Offices, which are usually placed under the Department of Planning and Statistics of the Ministries.

In the case of MAFF, the Office of Project Coordination and Monitoring and Evaluation (PCMEO) was established in 2004. The system is decentralized, giving all the authority to the implementing departments. The M&E Offices do not have legal authority to directly monitor and evaluate the outputs and outcomes of the activities and projects carried out by implementing departments. Hence, M&E activities are largely limited to the consolidation of reports. The institutional capacity of the M&E Offices is generally underdeveloped. Some constraints faced by implementing agencies include the limited number of staff with limited skills, and a lack of resources and authority.

The indicator system for M&E – In support of the current NSDP Monitoring Framework, a "two-tier structure" indicator system has been adopted. At the national level, the first tier, 43 core indicators have been set, in line with macro-development goals and the Cambodia Millennium Development Goals (CMDGs). At the line ministry/agency level, the second tier, sets of performance indicators have been developed based on the NSDP focus, CMDG indicators under its jurisdiction, and other indicators relevant for sectoral-level monitoring. A third tier of indicators may be added at the ministry/agency level to monitor programme and sub-programme activities.

The country-level development indicators for ARD Programmes – Cambodia's experience in using the indicator system as a tool for monitoring and evaluating ARD projects is still in its early stage. The institutional capacity and various underpinning infrastructures for an effective development indicator system are still underdeveloped. However, there have recently been significant steps taken to improve the system. Key milestones for the various attempts made to upgrade the system include: the enactment of the Statistics Law; the establishment of the National Statistical System (NSS) and the National Institute of Statistics (NIS) in MoP and the adoption of the Statistical Master Plan (SMP). These highlight the growing need for ample, timely, reliable and quality statistics relevant to development endeavours in the country. To date, notable improvements have been made in the areas of formal structure, management, staff training, dissemination practices and accessibility of data.

The current NSS is: (i) external funding-dependent and donor need-driven; (ii) fragmentary and disorganized, due to lack of agreement of statistical activities and standardized procedures; (iii) General Data Dissemination System (GDDS)-based; and (iv) largely decentralized. The first two features were reported to have imposed many limitations on the development process towards harmonizing official statistics in the country. This is due to a lack of or unstable financial support, which resulted in the piecemeal development of official statistics in the country. Data produced were largely aimed to meet the needs and priorities of external donor programmes, rather than the country's own perceived needs for relevant and appropriate data for monitoring national programmes. The lack

of consensus on priorities for statistical activities and standardized procedures were said to have caused difficulties with processing, analysis and interpretation. These resulted in a limited use of the data for policy, planning and programme formulation and evaluation.

The key sectors that make up the indicator system to provide economic, social, demographic and environmental statistics include agriculture, health, nutrition, education, commerce and the economy. A relatively large stock of indicators related to these sectors is available in the CAMInfo database produced by NIS of MoP. In addition, e-data of the Economic Institute of Cambodia (EIC), accessible via a prepaid card, is another online source of official statistics and indicators related to the country.

Statistics Law 2005 sets out a clear demarcation of responsibilities and relationships between ministries/agencies that are NSS stakeholders. Pursuant to the law, NIS is responsible for preparing official statistics policies, coordinating, and prioritizing activities, standards and methods necessary for creating an integrated NSS. Various ministries/agencies collect and produce statistics as part of their work. Some data come from administrative systems and others from statistical enquiry.

Based on the NSDP monitoring framework, 26 out of 43 core indicators are to be updated on an annual basis through the collection of administrative statistics. In general, indicators on macroeconomics, the labour force and employment, agriculture and food production, and education and literacy are suggested to be updated and disseminated annually. Most of the health and nutrition indicators are to be disseminated every two years; however, it was suggested that some of these should be disseminated annually.

Hard copy publications have traditionally been the main medium of dissemination for government statistics. To date, the usual hard copy publication known by users is the *Statistics Yearbook* published annually. Other forms of dissemination adopted by the NIS include: (i) Web sites; (ii) CD-ROMs (e.g. CAMInfo CD-ROMs); (iii) e-mail; (iv) the Data User Centre; and (v) the library. Necessary metadata on statistics series explaining the detailed methodologies used for the various statistical collections, periodicity, timeliness and dissemination are accessible on the GDDS Web site.

The lack of guidelines for setting national standards was cited as a major problem with much of the statistics work in Cambodia. The use of different methodologies has caused confusion and difficulties with data analysis and interpretation. For instance, data on income and poverty abound, yet poverty analysts were reported to have difficulties in drawing conclusions from these data. Moreover, there are concerns over the quality, timeliness and reliability of the data, especially those collected through the administrative system. Data gaps were also observed in some key areas such as economic statistics, finance, health, education and agriculture. The lack of financial and human resources has been cited as major constraints in efforts to develop NSS and overall official statistics.

The ARD framework – Results obtained from a series of consultations with a number of experts revealed general agreement on the usefulness of the proposed ***Sourcebook*** as a toolkit with a wide range of indicators that can be adapted/ adopted for ARD programmes. Access, use and satisfaction indicators were all felt to be relevant with respect to the policy, planning and M&E dimension.

The subsector indicators – Findings indicated that almost all indicators proposed in the *Sourcebook* are appropriate and feasible, although nearly half of the indicators were not yet available in the country. The agribusiness and markets, community-based rural development, rural finance and water resources management are the subsectors that have very few indicators proposed in the *Sourcebook* compared to other subsectors.

It is not advisable at the moment, however, to use the findings to draw conclusions on the adequacy or inadequacy of ARD indicators in the country. In fact, an expert in charge of the CAMInfo Unit in MoP confirmed that the current database contains more than 5 indicators, but they are mostly different from the proposed ones. This may not necessarily mean that the country experts have lagged behind in terms of the development and use of indicators; they may simply be different from the proposed ones. Should time permit, a more extensive review would surely provide an even clearer picture on the country-level indicators used in various subsectors.

Data supply for core indicators – Administrative records remain the main sources of data for at least 26 NSDP core indicators that should be collected and monitored on an annual basis. The rest of the core indicators, mostly outcome/ impact indicators, are to be supported by data supply from periodic and large surveys/censuses. Important periodic and large surveys/censuses conducted to date include agricultural surveys (e.g. crop cuttings, marketing surveys, and production cost surveys), demographic and health surveys, socio-economic surveys, inter-censal population surveys, child domestic worker surveys, child labour surveys, labour force surveys, industrial establishment surveys and the population census. The CAMInfo database and the *Statistics Yearbook* produced by MoP, and the e-data produced by the Economics Institute of Cambodia are important sources of data and official statistics for the national core indicators and the proposed ones.

To date, it is understood that Cambodia's capacity to supply data for core indicators is still limited, despite significant improvement made as a result of adopting the General Data Dissemination System, the Data Quality Assessment Framework and the integrated dissemination strategy. Data sources are still not adequate to meet the multiple needs of all relevant data users. Considering the context where technical, institutional and financial limitations still prevail, it is believed that there is still a long way to go before Cambodia could become fully

capable of building a system that produces and supplies adequate data for core indicators in line with the international standards.

Conclusions and recommendations – The study's findings suggested that Cambodia's experiences related to M&E, statistics and indicator systems are generally limited. Nevertheless, the road ahead is not an impossible journey. A better prospect for an improved capability of the country's M&E, supported adequately by timely and quality statistics inputs, is imminent, should the following recommendations be taken into consideration:

- The SMP roadmaps should be vigorously pursued.
- A systematic inventory of current indicators used within and outside the national institutions should be conducted.
- Harmonization and standardization of national M&E system should be proactively promoted.
- The M&E Units should be empowered with broadened legal authority and privileges.

The results of the study indicated an acceptance of the proposed Guidelines. In view of further improving the Guidelines, the following recommendations are made:

- Some indicators need to be transferred to appropriate subsectors, including indicators on livestock values/volumes, agricultural imports/exports and forest area.
- Some indicators of significant importance for Cambodia need to be added to the proposed Guidelines, including indicators on agribusiness and markets, community-based rural development, fisheries and aquaculture, forestry, livestock, and policy and strategy.
- Some indicators were considered neither appropriate nor feasible, so it was suggested to delete them from the Guidelines. These included indicators on ARD, agribusiness and markets, and water resources management.
- Modifications of indicators including the simplification of language or insertion/deletion of words used for constructing the indicators need to be made to improve clarity and understanding of indicators by users. It was suggested that some indicators be modified, including those on research and extension, agribusiness and markets, policy and strategy, rural finance and food security.
- The current global initiative to strengthen M&E and indicator systems from the conceptual to implementation level should be expanded. Capacity-building programmes in the areas of M&E and indicator systems development should be considered.

The Monitoring and Evaluation (M&E) information systems are designed within a specific institutional framework and according to its particular needs. They cater to the institutions, programmes and projects that they have to evaluate at different levels. Some systems are at the project level, but they are exceptions: they were not considered priorities at the moment of project development and tended to be substituted by the sectoral approach at the time of results-based management.

Basically, two levels, sectoral and the subsectoral, can be identified in the aim to implement monitoring arrangements based on the following indicators.

At the global level, the validity of the use of systems such as the Development Indicators National System (SINASID) depends on its use within a framework of global management by results. But since the country does not have an institutional planning system provided by law and equipped with the suitable technical apparatus for such an aim, there are real limitations to joint programming with the donors, which have continued with respect to the national systems of information in terms of evaluation by outcomes.

The concept of the sectoral M&E system known as the Follow-up and Evaluation System for Learning (SISEVA) was developed within the sectoral approach, together with the construction of a sectoral programme framework, the National Strategy for Productive Rural Development (ENDRP) – ProRural.

There are five components of ProRural. Three refer respectively to forestry, research and innovation, agribusiness and markets. A fourth refers to a combination of several items: rural development, community-based development, sustainable land and crop management, and rural finance. The fifth refers to basic infrastructure development, an item that is not part of the proposed list of indicators.

All national indicators can be found in the list of projects from the Rural Development Institute (IDR) in SISEVA, or in the evaluation frameworks of projects or isolated programmes. Follow-up therefore depends on the information flows from the institutions to SISEVA, which is limited to 30 indicators of early results and limited impacts. The operation of this system depends both on the structural conditions of the sector's institutions (the Agricultural and Rural Public Sector [SPAR]), which are not optimal for the effectiveness of the evaluation exercise; and on the demands of global planning, which are also seriously limited by the lack of a national planning system.

Success in the implementation of the Sector-Wide Approach Programme, as in ProRural, fundamentally depends on the institutional capacity of the sector being implemented. Implementation is a dynamic process that requires coordination, leadership, openness and motivation for change.

For these reasons, both the national and the sectoral level require additional institutional effort and more fluid relationships in both directions. The relationship between the sectoral and the national level is clear, since strategic outcomes of the former must be part of the national objectives.

One important point to mention is the actual restructuring process of ARD policy undertaken by the Nicaraguan Government. This process led to structural changes of the ProRural programme framework to create a new component for food security policy, as well as deep modifications in some of the current ones. These changes were known in the last trimester of 2007, i.e. after the completion of the country report.

Despite being too early to access the indicators due to their not having been reviewed to date, a study of the ARD proposal and the ARD indicators in Nicaragua was conducted using the current logical framework of the major projects and institutions related to rural development.

The key finding related to the data supply situation is that the statistical systems act independently from the evaluation systems, which are fed by institutional records, combined with their own studies and completed through the user surveys or household surveys.

In territorial or focused projects, many of which have already been concluded, one does not resort to national statistics, but rather to own records and ad hoc studies contracted by the project.

The Sectoral Statistical Systems such as that of the Ministry for Agriculture and Forestry (MagFor) serve as a database for National Accounts, but do not provide relevant information for the Ministry's management and planning.

The statistical system could be modified and adapted to the particular demand for analytic information generated by evaluation systems; in fact, its modification and reorganization has already begun, but it is not yet operational. According to the National Strategy of Statistical Development (ENDE), the National Statistical System (SEN) is weak and outdated, and therefore urgently in need of modernization and strengthening.

Finally, a significant aspect worth mentioning is the government's announcement, made in the Validation Seminar, that it intends to integrate this study in the conceptual organization of sectoral information for the National Strategy of Statistical Development being implemented in the country.

Nigeria has several policy documents that focus on poverty reduction and agriculture growth. These include: the National Economic Empowerment and Development Strategy (NEEDS) 2004 (federal and state versions), which provide an overarching strategy; the National Agricultural Policy (NAP 1988, 2001); the Rural Sector Strategy (RSS); and the Integrated Rural Development Policy Thrust (IRDP) 2004.

The government development strategy is to diversify the productive base of the economy away from the oil and gas sector, and to move towards market-oriented and private sector-driven economic development with strong local participation. Agriculture is seen as an instrument for poverty alleviation.

There are many agencies involved in M&E for ARD – both within the Ministry of Agriculture and externally. It is felt that greater coordination among agencies, leadership and standardization of procedures will enhance M&E results.

The organizations that were projected as possible candidates for leadership of M&E system are: the Plan Coordination Unit of MOA, the National Planning Commission, the National Bureau of Statistics, the Budget Office of Ministry of Finance and the National Poverty Alleviation Programme, among others.

The results of the surveys carried out by the NSO, particularly those relevant to the measurement of outcomes and impacts, are accessible to the M&E system, e.g. MICS, CWIQ and LSMS.

The World Bank, the African Development Bank (AfDB), the International Fund for Agricultural Development (IFAD) and the UK Department for International Development (DFID) are the leading donor agencies. The M&E system for donor-assisted projects tends to be more elaborate than the government-funded projects. The M&E in the entirely government-funded projects is limited to monitoring physical and financial targets.

Funding for the M&E work is an issue. A suggestion was made to make it obligatory to earmark a certain percentage of projects funds for M&E. It was suggested that providing a legal basis for M&E and constituting an independent commission for M&E, on the pattern of Auditor General Office with separate funding, will improve M&E.

M&E results are not commonly used by the Parliament, statesmen and senior officials for decision-making or for resource allocation. There is a need for building the technical capacity of personnel in M&E units in different line departments. In particular, the need was expressed for training in concepts such as the "logical framework".

The indicators on the list that were identified for reconsideration included: the US$1 poverty line, carbon sequestration, and increase in employment. It was suggested that an additional indicator, “quality of water in reservoirs”, be added to the list of core indicators. The access, use and satisfaction indicators were generally found useful.

Country study 4 – Senegal

This country study was considered relevant and timely for Senegalese counterparts as the government and partners are engaged in the process of strengthening and rationalizing the country's M&E system for more effectiveness, both at global level and the sector level. Several high-ranking government officials attended the two-day Validation Seminar and actively participated in the discussions.

Senegal, like most African countries, has prepared and adopted a Poverty Reduction Strategy (PRS) as the overall development framework. Given their importance in the economy, ARD subsectors are to contribute significantly to poverty reduction. Projects and programmes in the ARD subsectors are being implemented with a focus on poverty reduction and food security.

A Poverty Monitoring Unit is located in the Ministry of Economy and Finance (MEF), with focal points in line ministries. They work under a National Steering Committee and an Inter-Ministerial Orientation Council chaired by the Prime Minister. However, in parallel to this structure, line ministries have units in charge of studies and planning, with responsibilities for the monitoring, evaluation and statistics of all activities within their own ministries and also of the Medium-Term Expenditure Framework (Cadre de Dépenses Sectorielles à Moyen Terme, CDSMT). These CDSMTs are to some extent articulated within the PRSP. At present, the system seems to have overlapping roles and its functioning is not fully satisfactory. Also, the formulation of the ARD strategies and policy within the overall strategy is not systematically developed. The results of the M&E are not yet used as a basis for budget allocation, which reduces its impact on decision-making at the highest levels.

Furthermore, within the ARD sector, no single unit has the overall responsibility for M&E and statistics, since there are several ministries with their own units with little coordination among them (Agriculture, Livestock, Fisheries, Forestry, etc). As a consequence, there is a diversity of M&E systems and indicators in the sector, and the government and partners have undertaken actions towards their better coordination, standardization and harmonization within the sector. The process is also being mainstreamed with the reform of the NSS and the elaboration of the National Strategy for Development of Statistics (NSDS).

A set of indicators has been selected for monitoring the PRSP, and at the sector level, programmes and projects have logframes and indicators. The assessment of the core indicators proposed in the study with respect to the current situation reveals that a large number of the proposed indicators are relevant and overlap with the indicators selected for PRSP or at the sector level. Overall, out of the 100 indicators proposed in the study, 55 were compiled in Senegal, with censuses/surveys as data sources for 42 indicators. However, the situation varies from one subsector to another and

some of the indicators are neither relevant nor feasible in the country's context. For example, data related to rural finance is very fragmented and very few indicators are actually compiled. The same applies to Community Development Programmes, where the indicators proposed are considered not feasible in Senegal.

Finally, it should be noted that Senegal has undertaken a major reform of its NSS, with the creation of a semi-autonomous National Agency for Statistics and Demography (ANSD) at the core of the system, and the elaboration of a NSDS with sectoral components. This process is an opportunity to better align and rationalize the data and M&E system at the global and sectoral level. Both global and sector activities within NSDS are to be articulated and driven by data requirements for design, implementation, M&E of PRSP and sector development programmes.

The United Republic of Tanzania has invested a great of effort in defining a framework and mechanisms for an effective and efficient M&E system for tracking the results of its National Strategy for Growth and Reduction of Poverty (MUKUKUTA), which serves as overall development framework. This was done through dialogue and consultations between all stakeholders including the government and development partners. A global M&E structure is in place with a set of clearly defined and regularly monitored indicators and published annual reports. There is also a MUKUKUTA Monitoring Master Plan, which provides a basis for planning and implementing the main statistics operations through a corresponding basket funding.

At the sectoral level, the Tanzanian Government has adopted a sector-wide approach (SWAP) to development, and the agricultural sector development programme (ASDP) is the main tool for the central government for coordinating and monitoring agricultural development and for incorporating nation-wide reforms. The ASDP framework and content have been jointly developed by the four Agricultural Sector Lead Ministries (ASLMs) – the Ministry of Agriculture, Food Security and Cooperatives (MAFC), the Ministry of Industries, Trade and Marketing (MITM), the Ministry of Livestock and the Ministry of Water (MOW) – and the Prime Minister's Office–Regional Administration and Local Government (PO–RALG), in close consultation with other stakeholders. Under ASDP, an intensive consultation process with all stakeholders has resulted in defining a short and long list of indicators, which are being discussed for the monitoring and evaluation of the programme. In parallel to ASDP, there are still stand-alone projects being implemented in the agriculture and rural sector with their M&E systems. Ultimately, the government aims to have all projects converge to ASDP. Some donors contribute through basket funds, but others persist in traditional funding mechanisms. It is too early to judge how the sector-wide M&E system will work in practice, but all efforts are being made for adopting practical solutions.

An important policy orientation in the United Republic of Tanzania is the Decentralization by Devolution (D by D), in which local governments are being empowered with allocated resources. At this level, a Routine Data System (RDS), mainly using administrative sources, is being developed to complement data coming from censuses and surveys for the monitoring and evaluation of impact and outcome of programmes.

The comparison of the core indicators proposed in the *Sourcebook* against what is currently available shows that many of the indicators in the core menu of indicators do not correspond exactly to the specific project/programme indicators. However, they are similar or close proxies. Also, some indicators were excluded because of the difficulties, both technical and financial, in collecting data or

in compiling data to establish the indicator. Also, the process of formulating indicators is continuous, so that projects/programmes review and/or refine the indicators over time.

The results of the M&E system are highly appreciated by decision-makers, since they are increasingly used as a basis for discussions on budgetary allocations to ministries and local governments. The implication is a growing demand for data with high standards of quality, timeliness and regularity, which is becoming a challenge for the system. There are weaknesses in the system, including the limited capacity of decentralized structures, both for M&E and for basic statistics methodology, concepts and standards. Also, since censuses and surveys are a major data source, the timeliness of the results do not always correspond to the requirements of the M&E system. The high demand is putting great pressure on the National Bureau of Statistics (NBS), which has limited human resources capacity. Therefore, capacity building at all levels, particularly at the decentralized levels, appears to be critical for the effective functioning of the M&E system.

Part 2 – ARD indicators in use in each country

A common issue in all the workshops was that, even though there was a general consensus that the generic list of indicators was useful and collectable, less than one-third of them were actually available in any single country. The situation in each country is summarized in Table A2.1

Table A2.1 Summary of generic indicators currently available

Subsector	Total indicators	No. of generic indicators currently available				
		Cambodia	Nicaragua	Nigeria	Senegal	The United Republic of Tanzania
A. Core ARD sector indicators	28	8	7	9	8	3
B. Agribusiness and market development	13	2	4	4	3	3
C. Community-based rural development	9		2	4		2
D. Fisheries (aquaculture)	6	3	3	1	1	
E. Forestry	13	5	3	3	5	3
F. Livestock	8	5	5	7	6	2
G. Policies and institutions	18	6	11	11	7	6
H. Research and extension	7	4	3	4		
I. Rural finance	7		5	5		4
J. Sustainable land and crop management	9	6	6	5	2	
K. Water resource management	13	1	7	3	6	4
Total	131	40	56	56	38	27

From the original list of approximately 13 indicators, Nicaragua and Nigeria claim to be producing 56; Senegal, 38; Cambodia, 4; and the United Republic of Tanzania, 27. Each country also provided an additional list of proxy or similar indicators currently available. When compared with the generic list, it was apparent that the gap was actually not large and that many of the alternative or proxy indicators were in fact very close to or even the same as those on the generic list. Nevertheless, the weak capacity of NSSs is still a major constraint to the establishment of effective M&E procedures.

Table A2.2 ARD Indicators available in the five pilot countries

Sector/ subsector	Class	Indicator	List of available indicators in each test country				
			Cambodia	Nicaragua	Nigeria	Senegal	United Republic of Tanzania
A. Core ARD sector indicators							
	Longer term outcome	% change in proportion of rural population below US$1 per day and below national poverty line	√		√	√	√
	Early result	% change in cost of transportation of agricultural products					√
	Early result	% of the population employed, underemployed, unemployed	√	√	√		
	Longer term outcome	% of the population with access to safe/improved water	√	√	√	√	
	Longer term outcome	Annual growth of GDP per capita (%)	√		√	√	√
	Early result	Prevalence of underweight children under five years of age (%)	√	√	√	√	
	Early result	Proportion of malnourished population	√	√	√	√	
	Longer term outcome	Ratio (proportion) of arable land area to total land area (%)	√	√	√	√	
	Longer term outcome	Share of poorest quintile in national income or consumption	√	√	√	√	
	Longer term outcome	Value added in the agricultural sector per agricultural worker		√	√	√	
	Longer term outcome	% change in area under all major crops		√	√	√	
	Early result	% change in value of agricultural imports		√	√		
	Longer term outcome	% change in market share of cooperatives/public-owned enterprises					√
	Early result	% change in number of local businesses opportunities (over a set period)		√	√		
	Longer term outcome	% change in private sector investments in rural areas	√				
	Early result	% of population who consider that they are better off now than 12 months ago			√	√	
	Longer term outcome	Annual growth (%) of income from rural non-agricultural activities	√	√	√	√	
	Early result	Increased share of export price (urban consumer price) realized at the farm gate		√			√
	Longer term outcome	Proportion (or ratio) of total value of agricultural sector exports to total agricultural sector value added		√	√		√

Sector/ subsector	Class	Indicator	List of available indicators in each test country				
			Cambodia	Nicaragua	Nigeria	Senegal	United Republic of Tanzania
B. Agribusiness and Market Development							
	Early result	% change in (number, value, volume of activities) managed by agro-enterprises				√	√
	Early result	% of farmers who applied/purchased minimum package of inputs during the last season		√	√		
	Early result	% of targeted entrepreneurs with access to market information			√		√
	Early result	Proportion of (%) agro-enterprises adopting an improved / certified hygiene/food management system		√			
	Early result	Proportion of target farmers (by gender) who are members of producer organizations	√			√	
	Early result	Proportion of producer organizations capable of meeting the production and marketing needs of their members	√		√		
	Longer term outcome	% change in value of agricultural inputs (imported and local)		√			
	Longer term outcome	Well-functioning food safety surveillance, risk analysis, inspection and testing system		√	√	√	√
C. Community-based rural development							
	Early result	Indicators of access, use and satisfaction with community-based rural development services			√		
	Early result	% change in number of community associations exercising voting power in local government budget allocation processes		√	√		
	Early result	% of target communities that have had a community-based rural development project			√		√
	Early result	Proportion of POs/NGOs with functional internal system of checks and balances		√	√		
	Early result	% of completed projects still functioning after 3 years			√		√
D. Fisheries (aquaculture)							
	Longer term outcome	Annual growth or % change in the availability of fish/ production per capita	√	√	√	√	
	Longer term outcome	Annual growth or % change in value of production from aquaculture, by location (country, region, district, etc.)	√	√			

Sector/ subsector	Class	Indicator	List of available indicators in each test country				
			Cambodia	Nicaragua	Nigeria	Senegal	United Republic of Tanzania
E. Forestry							
	Early result	% increase in tax and royalty fees collected from the forest sector	√			√	
	Early result	Annual growth or % change in area under sustainable management (certified forest area, in ha)		√	√	√	√
	Early result	Proportion of forest area under private or communal ownership	√				
	Longer term outcome	% change in country's forested area	√	√	√		√
	Longer term outcome	% of targeted households benefiting from employment in the forest sector				√	
	Longer term outcome	Annual growth or % change in rural household income from the forest				√	√
	Longer term outcome	Rate of deforestation	√	√			
	Longer term outcome	Ratio of forested land area to total land area (%)	√		√	√	
F. Livestock							
	Early result	% of target farmers/herders (by gender) aware of improved breeds, feed, veterinary services and range management techniques	√	√	√		√
	Longer term outcome	% change in production/sales of animal products	√	√	√	√	
	Longer term outcome	% change in livestock values	√	√	√	√	
	Longer term outcome	% change in livestock numbers	√		√	√	
	Longer term outcome	Annual growth of animal population	√	√	√	√	√
	Longer term outcome	Livestock birth rate, by species, by area			√	√	
G. Policies and institutions							
	Early result	% change in number of local job opportunities over a set period			√		√
	Early result	Annual growth of food production (%)	√	√	√	√	√
	Longer term outcome	% change in value of agricultural exports		√	√	√	√
	Longer term outcome	Annual growth of income from the agricultural sector (%)	√	√	√	√	
	Longer term outcome	Proportion of land poor or landless population to total population (or agricultural population)	√	√		√	
	Longer term outcome	Ratio of average income of the richest quintile to the poorest quintile (%)	√	√	√		

Sector/ subsector	Class	Indicator	List of available indicators in each test country				
			Cambodia	Nicaragua	Nigeria	Senegal	United Republic of Tanzania
H. Research and extension							
	Early result	% change in number of smallholders (by gender) who use (apply, adopt) technology advice introduced by the extension system	√	√	√		
	Early result	% of farmers contacted by extension service in the last two weeks	√	√	√		
	Early result	Proportion of target farmers (by gender) providing input to agricultural research system	√	√	√		
	Longer term outcome	% change in yields resulting from use of improved practices	√		√		
I. Rural Finance							
	Early result	% change in number rural population accessing financial products for economic investments		√	√		√
	Early result	% or rural inhabitants using financial services		√	√		√
	Early result	Ratio of borrowers to savers		√	√		
	Longer term outcome	% change in access to formal credit		√	√		√
	Longer term outcome	% change in access to formal credit for women and minority groups		√	√		√
J. Sustainable land and crop management							
	Early result	Proportion of target farmers (women, men) who apply or have adopted sustainable crop production practices in their farms	√	√			
	Early result	Proportion of target farmers aware of sustainable crop production practices, technologies and inputs	√	√	√		
	Longer term outcome	% change in land access for women and minority groups	√	√			
	Longer term outcome	% change in revenues from natural resource use			√		
	Longer term outcome	% change in crop yield	√	√	√	√	
	Longer term outcome	% change in formal land transactions	√	√	√		
	Longer term outcome	% reduction of flood risks	√	√	√	√	

Sector/ subsector	Class	Indicator	List of available indicators in each test country				
			Cambodia	Nicaragua	Nigeria	Senegal	United Republic of Tanzania
K. Water resource management							
	Early result	% change in number or proportion of target farmers (by gender, tenure, head- and tail-enders) with access to a functioning (reliable, adequate) irrigation and drainage network			√	√	√
	Early result	% change in number or proportion of water users aware of roles and responsibilities of water users association members	√		√		√
	Early result	Proportion of service fees collection to total cost of sustainable water and irrigation activities and functions			√	√	
	Longer term outcome	% change in types of crops grown in all parts of the irrigation and drainage (I&D) system		√		√	
	Longer term outcome	% change in average downstream water flows over the project period during the dry season		√			
	Longer term outcome	% change in crop yields in all parts of the I&D system		√		√	√
	Longer term outcome	% change in cropping intensity in all parts of the I&D system		√		√	
	Longer term outcome	% change in GDP created by irrigated agriculture		√			
	Longer term outcome	% change in soil loss from project watersheds		√			
	Longer term outcome	% of irrigation schemes that are financially self-sufficient		√		√	√

Table A2.3 Alternative and substitute indicators used in the five test countries

Level	Proxies
A. Core ARD sector indicators	
	No. of products traded and publicized on markets, through the radio, leaflets, fairs and web pages
	% of farmers who receive technological assistance that have adopted the recommended practices
	Increase of equity among social groups with respect to food access
	Levels of food production, by category of foods
	Levels of food reserves
	Reduction of illness related to food intake habits
	Volume of crop production (other than rice)
B. Agribusiness and market development	
C. Community-based rural development	
	No. of organizations of youth groups and women who have access to direct financing
	% of women and girls in wage employment (agriculture, industry, services)
	Land tenure security index
	Land titles to farmers (% of total agricultural land)
D. Fisheries (aquaculture)	
	No. of municipal financing institutions that have started to diversify their offer of financial services and microcredit
	% of beneficiaries with access to credit fund who are women
	Credits up to pre-specified target level approved and disbursed
	Domestic credit
	Level of total arrears
	Net lending/net borrowing; saving
E. Forestry	
	Change in area covered by forest and woods
	Fuel wood dependency (% of households)
	% of households with access to common property resources
	% of employed persons in agriculture, hunting and fishing
F. Livestock	
	No proxy indicators were suggested for livestock

Level	Proxies
G. Policies and institutions	
	% improvement in human development and poverty indicators at the municipal level
	% of chronicle undernourishment in children under five years of age
	% of rural families served who have increased their ability to formulate training plans for employment and business
	Change in external trade balance with major partners
	Incidence of disease related to hygiene
	Increase of basic grains production in the Pacific, Central and Northern regions of Nicaragua
	Rural wage rate of unskilled labourers
	Total volume/value of agricultural exports by year
	Total volume/value of agricultural imports by year
	Yields and agricultural productivity
H. Research and extension	
	No. of technological themes disseminated
	% beneficiary groups that implement appropriate technologies for natural resources preservation
	% farms with implanted agroforestry with efficient practices of cattle feeding
I. Rural finance	
	No. of families receiving new financial products from local financial services providers
	No. of non-bank financial services providers strengthened through an institutional support programme
J. Sustainable land and crop management	
	% of rice cultivated area destroyed by drought and flood
	% of households affected by natural calamities
	% of small- and medium-scale farmers that use improved and environmentally friendly productive practices, including diversification
	Environmental quality index at the household level
	Land tenure security index
K. Water resource management	
	% of Farmer Water User Communities (FWUCs) with capacity to operate and maintain their I&D systems
	Irrigated area (% of rice area)

Annex 3: M&E capacity assessment scorecard

In order to facilitate the assessment of national M&E capacity, a checklist of questions to be addressed is provided, which may be used in two ways. The short method is appropriate when the primary objective is to raise awareness and stimulate interest in M&E capacity building in general. The full method is more suitable when the end objective is to prepare a proposal for an M&E capacity-building programme.

The short method. The short method is based on group discussions only and is suitable as a workshop exercise. The workshop consists of potentially interested stakeholders, possibly including donors and representatives of international organizations. Using this method, the full assessment could be completed in a few hours. It involves no data collection per se, but depends on having a well-informed group of stakeholders – including representatives from the private sector, civil society, and possibly donors – who are already familiar with ongoing M&E activities in the country. The discussants use the checklist of about 3 questions and their own knowledge of how M&E works in their country to ascertain a country score. The score represents a rough measure of the gap that separates the current, less-than-ideal situation from the ideal situation.

The full method. The full method is longer and involves data collection using surveys and interviews with a broad cross-section of data users and providers. The survey questionnaire should be built around the same checklist. This could be an appropriate assignment for a task force or consultant. The assignment would involve the design, implementation and initial analysis of the survey, including the preparation of a questionnaire to be administered to a carefully selected sample of users and providers. This phase could take several weeks.

Whichever route is used, the objective is to accumulate sufficient information to fill out the scorecard. The scorecard is divided into five sections: Basic (project-level) M&E capacity; sector-level M&E capacity; poverty monitoring; national-level M&E capacity; subnational-level M&E capacity; and potential for expansion. Each of the sections contains from five to eight questions that the focus groups are required to address. Next to each question are three multiple-choice answers. Basically, the groups are required to focus on different M&E activities and to rank the country capacity and experience in each area on a score of –3 (= no capacity;

1 = very limited capacity; 2 = some capacity; 3 = good capacity). The groups will review each question individually, mark the most appropriate answer and record the matching score in the box on the far right of each question. When all the questions have been completed, the scores are added up section by section, and the totals are then transferred to a summary scoresheet.

Table A3.1 M&E capacity assessment scoresheet

Summary Scoresheet	Scores	Max
A. Basic (project-level) M&E capacity		24
B. Sector-level M&E capacity		12
C. Poverty monitoring		15
D. National-level M&E capacity		21
E. Subnational-level M&E capacity		15
F. Potential for expansion		9
TOTAL		**96**

Since the answers are obviously subjective, they cannot be interpreted in absolute terms, but in general terms only. Countries with an overall score of less than 25 points usually have the least capacity; countries with 25–50 points have some fairly limited M&E activities; and those with 50–75 points have strong competencies. Countries scoring over 75 points are considered to have strong overall capacity. Having ascertained the country's overall capacity level, the discussants should then refer back to the questions on a section-by-section basis to identify where specifically capacity is weakest.

M&E capacity assessment scorecard

1. Basic (project-level) M&E capacity					
Most ARD programme/projects have an active M&E component.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Most ARD projects have their own M&E units.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
The logframe is generally used for project design and M&E.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
The monitoring of inputs and outputs is generally well executed.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Most projects use computerized Management Information Systems (MISs).	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Most projects produce regular monitoring reports.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Monitoring reports influence the allocation of resources for the next reporting period.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Project M&E units have the capacity to carry out surveys on intended beneficiaries.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Project-level M&E capacity score =					
2. Sector-level M&E Capacity					
Project M&E activities are well coordinated.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Sector ministries concerned with ARD have M&E units.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
These units produce timely, reliable and useful progress reports.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
These units work on both the monitoring of performance and the measurement of results.	[1] Strongly disagree	[2] Disagree	[3] Neither agree nor disagree	[4] Agree	[5] Strongly agree
Sector-level M&E capacity score =					

3. Poverty monitoring (Which of the following elements feature in the poverty monitoring programme?)				
Is there a Poverty Monitoring Unit and how effective is it?	[1] No unit	[2] Not very effective	[3] Effective	[4] Very effective
Is there a National Household Survey (including household consumption data) executed every 3-5 years by the NSO or equivalent?	[1] No survey	[2] Unreliable or outdated survey	[3] Adequate survey	[4] Good survey
Is there regular collection of service delivery indicators?	[1] No collection	[2] Unreliable collection	[3] Fairly good collection	[4] Good collection
Are there qualitative poverty surveys/studies?	[1] No studies	[2] Poor studies	[3] Fairly good studies	[4] Good studies
Are there annual/biennial poverty monitoring reports	[1] No Reports	[2] Irregular reports	[3] Fairly good reports	[4] Good reliable reports
Poverty monitoring capacity score =				
4. National-level M&E capacity				
Is there a National M&E Unit and how effective is it?	[1] Non-existent	[2] Very limited influence	[3] Moderately effective	[4] Very effective
Is there a National M&E Coordinating Committee (or equivalent) and how effective is it?	[1] Non-existent	[2] Very limited influence	[3] Moderately effective	[4] Very effective
Does the M&E system produce regular (annual) PRS progress reports?	[1] No Report	[2] Irregular Reports	[3] Fairly good reports	[4] Good reliable reports
Are there econometric modelling studies?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
Does the M&E system have the capacity to undertake impact evaluation studies?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
What capacity is available to plan and execute a programme of household surveys?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
How easy is it for interested users to gain access to primary data sets for carrying out further research and analysis?	[1] Not Possible	[2] Very difficult	[3] Moderately difficult	[4] Moderately easy
National M&E capacity score =				
5. Subnational-level M&E capacity				
Are there the necessary structures at the subnational level to carry out M&E activities?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
Are regular (annual) PRS progress reports produced at the subnational levels?	[1] No Reports	[2] Irregular reports	[3] Fairly good reports	[4] Good reliable reports
Is there a standard financial record-keeping and accounting system?	[1] No system	[2] Very limited system	[3] Some capacity	[4] Good system
What capacity is available at the subnational levels to produce annual estimates of agricultural and livestock production?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
What capacity is available to carry out household surveys?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
Subnational M&E capacity score =				
6. Potential for expansion				
Is there any experience and/or capacity for community-level monitoring?	[1] No capacity	[2] Very limited capacity	[3] Some capacity	[4] Good capacity
Do the M&E activities include any form of corruption monitoring?	[1] No	[2] Very limited	[3] Some	[4] A lot
Does the media (radio, newspapers, etc.) promote a wider dissemination and discussion of M&E results?	[1] No	[2] Very limited	[3] Some	[4] A lot
Future directions score =				

The Sourcebook provides a number of workable approaches for designing an M&E system that would be of greatest relevance to different agricultural and rural development (ARD) activities, projects and programmes, and degree of data availability.

A set of 19 priority indicators based on the criteria of comparability, availability and relevance have been identified for the purpose of international comparisons. It is expected that most countries, regardless of the stage of development of their monitoring system and statistical capacity, will be in position to provide periodic data on these indicators.

A comprehensive set of 86 indicators validated in countries in Asia, Africa and Latin America, covering all subsectors of ARD and some thematic areas, offers M&E professionals, project planners and policy-makers a ready-made menu to select the indicators that best suit their needs. In preparing the menu, due care has been taken to include indicators that are workable even in situations where data availability is less-than-ideal.

ISBN 978-92-5-106082-7

9 789251 060827

TC/M/I0380E/1/10.08/3000