
PART II

SOME COUNTRY EXPERIENCE

Part II focuses on the experience of several countries that have successfully built well-functioning M&E systems, and on Africa. A common question asked by skeptical government ministers and officials is whether there are actually any countries that have successfully created a national M&E system. The answer to this question is an unambiguous “yes.” Many developed, and a small but growing number of developing, countries have succeeded in building well-performing, whole-of-government M&E systems. Chapter 5 outlines what a “successful” government M&E system looks like; three dimensions of success are proposed. Three such countries—Chile, Colombia, and Australia—are presented in chapters 6, 7, and 8. The special case of M&E in Africa, where there are many severe capacity constraints, is considered in chapter 9.



5

Good Practice Countries— What Does “Success” Look Like?

Officials in many developing countries are interested in the experiences of governments with well-functioning M&E systems. Such governments have accumulated valuable lessons about how to set up and successfully manage an M&E system—what to do, how to do it, and the pitfalls to avoid. These governments also showcase the cost-effectiveness of M&E.

As noted in chapter 3, there are dangers in attempting to uncritically replicate another country’s model. In Latin America, for example, many countries look to the case of Chile—which has a very strong and disciplined M&E system—and would like to apply it to their own country. However, Chile possesses a specific and rare combination of characteristics: a highly centralized budget system, a highly capable and extremely powerful finance ministry, sector ministries and agencies that closely follow the rules and procedures set down by the finance ministry, a disciplined civil service, and a highly capable academic community. These are all success factors for the Chilean government’s M&E system. But there is only one Chile.

This is not to say that Chile’s experience, or that of other countries with successful M&E systems, is irrelevant to other countries, even those with relatively weak public administrations. The very process of comparing any individual country with another that possesses a successful M&E system is illuminating; it helps reveal the reasons for that success, and it clarifies how easy or difficult it might be to replicate that success.

The concept of a successful M&E system also requires some examination. The definition applied throughout this volume is as follows:

The successful institutionalization of M&E involves the creation of a sustainable, well-functioning M&E system within a government, where good quality M&E information is used intensively.

Three dimensions of success are stressed here:

1. **Utilization of M&E information.** The information is used in one or more of the four principal ways outlined in chapter 3, that is, to support government policy making, including performance budgeting or national planning; for policy development and analysis and program development; for program and project management; or for accountability purposes. Utilization of M&E information can, of course, range along a spectrum from zero or negligible to substantial (or intensive). Intensive utilization can be viewed as reflecting the mainstreaming of the M&E

function in the government. Most evaluators and evaluation offices in governments and donor organizations have a surprisingly poor understanding of the extent to which their evaluation findings are or are not used by others.

2. **Good quality M&E information.** Governments differ considerably in terms of what they conduct under the heading of “M&E.” Some stress a system of performance indicators—focused on national development goals; ministry goals; and lower levels of ministry outputs, service delivery, and processes. Others focus on carrying out various types of evaluation, such as rapid reviews, rigorous impact evaluations, or other types and methods of evaluation. There are standards against which M&E can be compared to determine if it represents good quality or not.¹ Most evaluation offices have some sort of quality control mechanism in place. Most, however, do not appear to conduct or commission formal evaluations of the quality of their M&E work.
3. **Sustainability.** This relates to the likelihood that the M&E system will survive a change in administration or in government ministers or top officials. Where the utilization of M&E information is firmly embedded—that is, mainstreamed—in core government processes such as the budget cycle, it can be said to be institutionalized and thus is likely to be sustained over time. Conversely, where M&E has only a handful of key supporters or is little used, or if it is largely funded by donors rather than by the government itself, then sustainability would be seen as less likely.

Three case studies of countries with good practice government M&E systems are presented in the next chapters: Chile, Colombia, and Australia. In-depth reviews of their M&E systems are available (Rojas and others 2005; Mackay and others 2007; Mackay 2004). None of the three can be considered to constitute a “perfect” M&E system in terms of the three dimensions of success outlined above. Each has strengths and weaknesses, as is evident in the following discussions.



6

Chile

The government of Chile has progressively developed its M&E system over a number of years, with most of the development having occurred since 1994. The system has been largely designed, implemented, and managed by the powerful Ministry of Finance (MoF), with the overall objective of improving the quality of public spending.

The system's development has also been influenced by fiscal pressures and the need to rein in overall government spending. Another influence has been the changing landscape of public sector reforms. In this context, the system has—appropriately—been developed in an opportunistic manner.

Main Components of the M&E System

The system has six main components. The first, long-standing component is the ex ante cost-benefit analysis of all investment projects. This was first introduced in 1974. This work is undertaken by the planning ministry. All the other components of the government's M&E system are based in the MoF. As noted in chapter 5, Chile has a very powerful and capable MoF. It plays a dominant role in the annual budget process and in the M&E system; the MoF is significantly more powerful than the sector ministries and agencies.

The second component of Chile's M&E system is performance indicators, which were first piloted in 1994 (see box 6.1). The MoF currently collects about 1,550 performance indicators, for all sectors.

The third component—comprehensive management reports—was introduced in 1996. These reports are prepared annually by ministries and agencies and report on their objectives, spending, and performance.

The fourth component is the evaluations of government programs, initiated in 1996. These follow a standardized format and comprise rapid reviews, which include a logframe analysis of a program, a desk review, and an analysis of existing data. Rigorous impact evaluations are the fifth component. They were introduced in 2001, entailing primary data collection and analysis usually based on sophisticated statistical techniques.

The sixth and most recently introduced M&E component is comprehensive spending reviews (2002). These reviews analyze all programs within a particular functional area and address issues such as inefficiency and duplication of programs.

A commendable feature of Chile's system is its “graduated approach” to M&E. It regularly collects performance information for all programs and

Box 6.1: Chile's Whole-of-Government M&E System

The six main components of the M&E system are as follows:

Ex ante cost-benefit analysis is required for all government projects (since 1974). This component is managed by the ministry of planning; it is the only component not managed by the MoF.

Performance indicators are collected for all government programs. They were first introduced on a pilot basis in 1994. The number of performance indicators has increased rapidly in recent years, from 275 in 2001 to about 1,550 currently. Of these, 25 percent relate to process issues, 57 percent to government outputs (that is, goods and services produced), and 18 percent to outcomes. Each ministry and agency provides the information to the MoF; there are about 11 indicators per entity. Entities are expected to have management information systems in place to produce this information. The MoF undertakes some data checking and data audits, and it includes the performance information in the budget bills it prepares each year.

Comprehensive management reports (1996). Each ministry and agency prepares one of these reports annually, based on MoF guidelines. The reports are intended to be the main public disclosure document. They report spending, use of funds, and performance; the reports thus draw heavily on the performance information that entities are required to produce and on the evaluations commissioned by the MoF. The reports also describe the progress made by the entity in achieving the formal institutional

commitments it has agreed to with the MoF; these comprise specific actions the entity has promised to implement to improve its performance (discussed in the main text). The draft reports are reviewed by the MoF and the ministry of the presidency, and entities make any necessary revisions. The final versions of the reports are sent to the Congress.

Evaluations of government programs (1996). Some 185 of these rapid reviews have been conducted so far (that is, until the end of 2006). They entail the clarification and agreement (between the MoF and the ministry or agency whose program is being evaluated) of detailed program objectives. A logframe analysis and desk review of program performance is conducted, drawing on available performance information. Their average cost is about \$11,000, and they usually take four to six months to complete.

Rigorous impact evaluations (2001). These evaluations involve primary data collection, sophisticated data analysis, and often the use of control groups. Eighteen have been completed so far, at an average cost of \$88,000 and taking up to 18 months to finish. Excluding defense spending and income transfer payments, more than 60 percent of government spending has been evaluated so far.

Comprehensive spending reviews (2002). These reviews assess all programs within a particular functional area or ministry. They examine issues of inefficiency and duplication of programs. Eight of these desk reviews have been conducted so far, at an average cost of \$48,000.

activities; these are also used in its evaluations. The evaluations of government programs are conducted selectively, particularly for newer programs. More in-depth impact evaluations are conducted later in the life of a program, especially where the impact is unknown or where the program is a government priority.

Performance Budgeting

A number of features of Chile's M&E system are important for its operation; they determine how much the M&E information the system produces is used. One feature is the close link between the system and the annual budget cycle. A danger

with any M&E system—whether it is a whole-of-government system or a sectoral system—is that it is managed by a stand-alone, specialist unit that operates separately from the mainstream activities of its host ministry. In Chile, the five components of the M&E system that are the responsibility of the MoF are managed by its management control division. This division works closely with the director of the budget, to whom it reports—and who has ministerial rank and is a member of the Cabinet—and also with the budget sections that have responsibility for overseeing the finances and performance of all sector ministries and agencies.

Indications of poor program performance are used in Chile as one trigger to warrant a more in-depth investigation of the causes, through a formal evaluation: either a rapid evaluation or a sophisticated impact evaluation. The MoF's budget directorate plays the main role in identifying government programs that should be evaluated. In preparing this evaluation agenda, the MoF endeavors to anticipate the information needs of the coming budget.

The agenda is also discussed with the ministries of the president and planning—indeed, these three central ministries comprise an interministerial committee that oversees all evaluations—and with the Congress, but it is clear that the main player is the MoF. The budget section heads in the MoF are also required to provide detailed comments on evaluation reports relating to the entities they oversee, and the evaluations are then discussed with MoF's budget director. Decisions concerning budget allocations may be taken at this stage.

During the budget process, the budget director meets with staff of the management control division and the budget sections to discuss each entity's budget proposals and the entity's overall performance. These meetings discuss the comprehensive management reports that each entity is required to provide¹—they include the entity's objectives, financial and performance information, evaluation findings, and progress against the performance targets² that were set during previous budget rounds.

This M&E information constitutes an important input into budget decision making, but it is only one input among others. As noted in chapter 3, there is rarely any direct formulaic relationship between good or bad entity performance and budget allocations. Thus, in some cases, poor performance by an entity will lead to reduced budget funding or even the termination of the program. But in other cases, poor performance of a program considered a government priority might necessitate a short-run increase in budget funding to correct the problems identified.

Management Improvement

A notable feature of Chile is the way the MoF uses M&E information to improve the performance of ministries and agencies. It does this in two ways. First, the performance indicators provide baselines of program performance, and the MoF agrees on performance targets for the coming budget year with each organization. The MoF monitors the extent to which the targets are met. In 2003, for example, about three-quarters of these targets were met.³ Second, when the MoF considers the recommendations made by the evaluations it has commissioned, it discusses them with the evaluated organizations and formally agrees on changes to the programs. In effect, the MoF imposes these agreements—known as formal commitments—on the organizations. It is quite unique for a MoF to systematically impose such management changes on ministries and agencies. Chile's MoF is able to do this because of its powerful role within the government.

Table 6.1 shows how the MoF uses the evaluations it has commissioned. These uses include minor changes to program management (for 24 percent of evaluated programs); major changes (38 percent); substantial program or organizational redesign (25 percent); institutional relocation of the program (5 percent); and program termination (8 percent).

Strengths and Weaknesses of Chile's M&E System

Chile's system can be assessed against the three criteria of a successful M&E system: high utilization, good quality M&E, and sustainability. As noted, M&E information is used intensively in budget analysis and decision making. It is also used intensively to impose program improvements on ministries and agencies. Last, it is used in reporting government performance to the Congress and to civil society.

What is missing from this list, however, is utilization of the M&E information by ministries and agencies themselves (other than those program changes that the MoF imposes on them). A recent World Bank evaluation of Chile's M&E system

Table 6.1: Utilization of Government Evaluations—2000–05^a

Effect on program	Minor adjustment of program, for example, improved processes or information systems	Major adjustment of management processes, for example, changed targeting criteria, or new information systems	Substantial redesign of program or organizational structure	Institutional relocation of program	Program termination
Percent programs affected	24	38	25	5	8

Source: Guzman 2007.

a. Percentages relate to evaluated programs.

found that utilization was low (Rojas and others 2005).⁴ The main reason for this is the low level of ownership—or acceptance—of the findings of the evaluations commissioned by the MoF. This weakness of Chile’s M&E system arises from its centrally driven, force-fed nature.

Another limitation of Chile’s centralized system is the apparent absence of incentives for ministries and agencies to conduct their own evaluations.

The available evidence on the quality of monitoring information and evaluations conducted by the M&E system indicates that the quality is broadly adequate and that the M&E information is thus broadly reliable—but no better than that. Although data verification and some data audits are conducted, there is no systematic approach to undertaking data audits of performance information.

The MoF contracts the evaluations out to academia and consultants to help ensure that they are conducted in an independent manner. Standardized ToRs are used for the evaluations, and this helps achieve some commonality in the issues evaluated and methods used; however, lack of data is sometimes a constraint on the quality of evaluations and the evaluation methods used. Evaluations are conducted within tight time constraints to ensure they can feed into the MoF’s budget analysis and decision making. The MoF also ensures that the cost of evaluations is kept as low as possible. These tight time and cost constraints sometimes result in an inability to collect

the primary data necessary to conduct rigorous impact evaluations, for example.⁵

The likely sustainability of Chile’s M&E system appears high. A series of budget directors in the MoF has worked to progressively develop the M&E system as an integral part of the budget process, and M&E is now embedded as a core function of the MoF. Given the preeminent role of the MoF, there do not appear to be any trends or pressures that would reduce the priority it gives to M&E. One weakness, already noted, is the low level of ownership of M&E by sector ministries and agencies. This constitutes, in part, a lost opportunity for the government as a whole to use M&E information more intensively; it also constitutes a lost opportunity to achieve a stronger performance culture within these sector entities.

The sustainability of the M&E system thus relies on the MoF’s continuing willingness and ability to drive the system centrally. If the government should ever decide to reduce the central power of the MoF, then the strength and utilization of the MoF’s M&E system might be called into question.

A more detailed list of the strengths and weaknesses of Chile’s system is presented in table 6.2.

Conclusions

The government of Chile has succeeded in creating a system whose monitoring information and evaluation findings are utilized intensively, particularly during the budget process. It has also suc-

Table 6.2: Strengths and Weaknesses of Chile’s M&E System

Strengths	Weaknesses
<ul style="list-style-type: none"> • “Graduated” approach to M&E. • Evaluations are conducted externally, in a fully transparent process, and are considered highly credible by other ministries and the Congress. • All M&E information is reported publicly and sent to the Congress. • The M&E system is closely linked to the information needs of the MoF, especially for the budget process. There is high utilization of M&E information in the budget. • Performance information is used to set performance targets for ministries and agencies; these are largely met. • The MoF uses evaluation findings to impose management changes on ministries and agencies. • The MoF closely monitors the extent of utilization of its evaluation findings. 	<ul style="list-style-type: none"> • Unevenness exists in the quality of evaluations conducted; this is probably caused by cost and time constraints imposed by the MoF. • Chile is probably not spending enough on evaluations. • There is a low level of utilization—because of low ownership—of evaluation findings by sector ministries. • There is an apparent absence of incentives for ministries and agencies to conduct their own evaluations.

cessfully driven management improvements in sector ministries and agencies. The M&E system includes what is in effect an “evaluation factory” for planning, commissioning, managing, reporting, and using the evaluations. This approach keeps the cost of evaluations, and thus the overall cost of the evaluation system, low: the finance ministry annually spends a total of around \$0.75 million on the M&E system,⁶ which is a very modest figure compared with the total government budget of \$20 billion.

This high utilization of M&E information is very impressive, and this alone makes Chile’s M&E system the strongest in Latin America and one of the strongest in the world. Chile has demonstrated that a whole-of-government M&E system can be built and operated at a relatively low cost. Many other governments in the Region are looking to emulate Chile’s M&E system, although a number of the success factors in Chile appear to be unique—such as the very powerful role of the MoF and the compliant nature of sector ministries and agencies.

Chile’s M&E system appears to be highly cost-effective: the MoF extracts considerable use from

the M&E information produced and does so at relatively low cost. One issue for the government to consider is whether it is spending enough on its M&E system. Sophisticated, wide-ranging evaluations can cost much more than Chile has spent on any of its evaluations in the past.

In Mexico, for example, a series of rigorous impact evaluations of the *Progresas* conditional cash transfer program in the 1990s cost several million dollars—compared with total government spending on the program of \$780 million in 1999. These evaluations thus constituted only a small fraction of total government spending on the program. The evaluations have been highly influential in persuading successive governments not only to retain the program but to scale it up significantly; by 2005 the government was spending about \$6 billion on the program (renamed *Oportunidades*) annually, covering 21 million beneficiaries, or about one-fifth of the Mexican population.

A question for the Chilean government to consider is whether it would ever spend a large amount on an evaluation of one of its programs, even if the program were one of its major priorities. Failure to spend adequately on evaluations will limit the

depth and reliability of evaluation findings, and this has probably been the cause of the quality problems of some of the evaluations commissioned by the MoF.

Finally, sector ministries' and agencies' low level of utilization of the system's evaluation findings constitutes an unexploited opportunity for the government. One way for the MoF to achieve

greater ownership of evaluation findings and to encourage the development of a performance culture in the civil service as a whole would be to pursue a somewhat less centralized, more collegial approach to the planning and oversight of the evaluations conducted. A broader base of support for M&E within the government would also further increase the likely sustainability of Chile's M&E system.



7

Colombia

The government of Colombia decided in 1991 to create an M&E system, which would be based initially on the World Bank’s own approach to evaluation. A constitutional requirement for evaluation was introduced later that year, and the Bank and other donors quickly followed with a range of technical and financial support. The government also introduced a series of laws, decrees, and regulations to buttress the M&E system.

During its early years, the main emphasis of the system was on monitoring information rather than evaluation. The system has waxed and waned, however, and by 2000 Colombia even considered abolishing the system—because of the perception of difficulties with the system’s management, and also because of doubts about its relevance to the government’s public sector reform agenda.

A substantial change of fortune occurred with the election of a reformist president in 2002. The president had been dismayed to note that the large increases in government spending in areas such as schools and health care had not been matched by corresponding increased performance (outputs and outcomes) in these areas. At the same time, he was strongly committed to a new culture of public administration based on social accountability—or social control. The president recommitted the government to a rejuvenated M&E system, and this was followed by a fresh infusion of donor support. These steps quickly strengthened the system.

Main Components of the M&E System

There are two main components of the system—which is known by its Spanish acronym, SINERGIA (Sistema Nacional de Evaluación de Resultados de la Gestión Pública, or National System for Evaluation of Public Sector Performance)—as well as several other components that have been or are currently being piloted. The most visible and most heavily utilized component is the subsystem for monitoring progress against all 320 presidential goals and the country’s development goals (as contained in the national development plan).

This subsystem (Sistema de Programación y Gestión por Objetivos y Resultados [SIGOB], or System of Programming and Management by Objectives and Results) records the goals, their strategies, baseline and target performance, and amounts spent by the government on them. Goal managers are also required to provide detailed explanations when goals are not met. All this information, including the contact details for each goal manager, is publicly available on a government Web

Box 7.1: Colombia's Whole-of-Government M&E System

The main components of the M&E system are as follows:

There are about 500 performance indicators relating to 320 presidential goals, and for each indicator, SIGOB records the objective, the strategies for achieving it, baseline performance, annual targets, actual performance against these targets, and imputed amounts spent by the government. Thus, SIGOB includes a large number of indicators on government performance. The information is disaggregated by region, including for the major cities.

In addition, where a target has not been met, the goal manager is required to prepare an explanation of the reasons for the shortfall. These exception reports are included in the SIGOB database, the core of which is publicly available on a real-time basis. The Web site also encourages accountability by identifying the goal manager and his or her ministry and formal position and e-mail address. It is the responsibility of ministries and agencies to supply the SIGOB data to the system's manager—the Department of Na-

tional Planning (DNP)—which undertakes some data checking. However, there is no formal system of data audits, and some concerns have been raised about ministries gaming (that is, distorting) the data they provide.

About 15 rigorous impact evaluations and institutional and management evaluations are under way, and another 22 are planned. Their cost ranges from \$15,000 to \$2 million; most cost between \$50,000 and \$200,000.

These evaluations are contracted out to academia or consultants, with oversight by the planning department in close collaboration with both the evaluated agency and the donors funding the evaluation. Rapid evaluations are also being piloted, with a view to mainstreaming them in the budget and planning work of the finance and planning ministries. The rapid evaluation methodologies have drawn on Chile's approach and on the rating approach of the United States' PART system (see boxes 6.1 and 3.2). The pilots currently being conducted are expected to cost between \$15,000 and \$25,000 each.

site (<http://www.sigob.gov.co/ini>). The President uses the SIGOB information intensively for political and social control.¹

SINERGIA's second main component is the series of evaluations it is conducting (see box 7.1). At the end of 2006, 15 evaluations were being conducted or had recently been completed, with another 22 planned for the following five years. The total cost of these evaluations is \$11.1 million. Unlike in Chile, no standardized types of evaluation are conducted; Chile has standardized ToRs, evaluation approach, and cost limits for each type of evaluation the MoF commissions. The three main types of evaluation in Colombia are rigorous impact evaluations, "institutional" evaluations, and "management" evaluations—the latter two focus on management and process issues. The most expensive (\$2 million) is a rigorous impact evaluation of a conditional cash transfer program.

These evaluations are collaborative exercises involving the planning department, the sector ministry or agency responsible for the program being evaluated, and donors. Most of the funding for these evaluations is provided through donor loans.

Most of the programs being evaluated are managed by the ministry of social protection, the family welfare institute, or the ministry of education.

A fourth type of evaluation is currently being piloted—rapid evaluations of those government programs that are either high priority or that have some suspected performance problems. The pilots are being conducted by the finance ministry and the planning ministry.² The intention is that rapid evaluations will eventually be mainstreamed in the core budget analysis and decision-making work of the two ministries. (In Colombia, there is a split budget: The finance ministry is responsible for recurrent spending; the planning ministry is responsible for the investment side, including education spending and cash transfers.)

The M&E system has other components, but they are weaker and much less fully institutionalized. One is an effort to partner with civil society—such as establishing consortia of NGOs and the media—in analyzing government performance. Another component is support for two municipalities that are undertaking performance monitoring and performance budgeting.

One final component is the preparation of performance budget reports to the Congress. These have been prepared by the planning ministry as an annex to the conventional budget documents. They present government spending on a programmatic basis and report the available M&E information on the performance of these programs. As noted in box 3.1, such performance reporting is a form of performance budgeting, albeit a weak type. It is particularly weak in Colombia because the Congress plays a weak role in the budget process, it has little technical support to enable congressmen to easily digest performance information and evaluation findings, and Congress' discussion of the annual budget tends to focus on narrower political issues (see annex B).

Accountability—Political and Social Control

A unique feature of Colombia has been President Uribe's strong commitment to the use of M&E information to enhance political control of the executive government and to support social control. The SIGOB database is loaded in his personal computer, and he uses this information in his monthly management control meetings with each minister and the DNP. During these meetings, the progress being made against each presidential goal is reviewed, and ministers are required to provide reasons for any shortfalls in performance. Performance indicators and ac-

tions to meet these targets are also agreed. The president uses this SIGOB information in his weekly town hall meetings in different municipalities around the country and also in an annual television presentation to citizens, in which he and his ministers discuss the government's performance and answer citizen questions on these issues.

This strong presidential commitment to using M&E information to monitor and report on the government's performance appears to be unique in Latin America—and perhaps in the world. It sends powerful signals to individual ministers and civil servants. However, there does not appear to have been a widespread adoption of M&E practices by ministries and subnational governments. Several ministries and agencies in the social sectors, as well as two municipalities, are currently working to strengthen their M&E systems with assistance from the DNP and donors.

Use of Evaluations to Support Government Decision Making

By early 2007 only three evaluations had been completed, so the opportunity to make use of evaluation findings to support government decision making in both the national budget and for national planning has been limited. But one notable example of an influential evaluation does exist (see box 7.2).

Box 7.2: An Influential Evaluation in Colombia

Familias en Acción is a government conditional cash transfer program that provides income support to poor families that commit to ensuring that their children receive preventive health care, enroll in school, and attend classes. The program was created in 1999 in response to an economic crisis.

A rigorous impact evaluation of the program was contracted out to external consultants, under the supervision of the planning department. The evaluation found that the program achieved impressive nutrition, education, and health impacts. These findings persuaded the government of President Uribe, who was elected

in 2002, not only to retain the previous government's program but to double its coverage, from 500,000 to 1 million poor families. In late 2006, the government decided to further increase the program's coverage, to 1.5 million families.

The *Familias en Acción* evaluation has cost \$1.5 million so far. Although this is a large amount, it is relatively small when compared with total government spending on that program (around \$100 million at the time of the evaluation). Because of its major influence on the government, this evaluation can be judged to have been highly cost-effective.

Source: Mackay and others 2007.

Strengths and Weaknesses of Colombia's M&E System

Colombia's system can also be assessed against the three criteria of a successful M&E system: high utilization, good quality M&E, and sustainability. The evidence base for such an assessment for Colombia is not as robust as it is for Chile, but some conclusions can nevertheless be drawn. Utilization by the president of the monitoring component (SIGOB) of the M&E system is notably high, for purposes of social and political control—that is, for accountability.

The quality of the monitoring data is unclear, but there are fears—held by some senior civil servants and by some influential members of civil society—that the data are not wholly reliable and that some of the information provided by sector ministries and agencies might be self-serving. The sustainability of this monitoring component seems highly likely, however, even after a change of administration and a new president. This is because of the usefulness of such data for the DNP and the president's office in their oversight of government performance.

The other main component of Colombia's M&E system is the ambitious program of rigorous impact evaluations, institutional evaluations, and management evaluations. Only a few evaluations have been completed, but one of them—*Familias en Acción*, one of the government's most important programs—has been highly influential and can be judged as having been very cost-effective. This is noteworthy because even a single, high-profile evaluation that influences a government's decisions can also influence more widespread acceptance of M&E by demonstrating its value.

The quality of Colombia's evaluations has not been formally reviewed, although there is some comfort concerning their likely quality and reliability: they have been contracted out to eminent academics, including some internationally renowned evaluators, and have had significant involvement from evaluators from the World Bank and other donors. The sustainability of the evaluation program is not entirely assured, however.

In the short term—that is, for the next five years—the evaluation agenda will be supported by a new World Bank loan; one feature of this agenda is a declining level of donor financial support for these evaluations, in the explicit expectation that the government will take up the slack by using its own budget funds to pay for them.³ This is likely to happen if the government judges the evaluations produced as worth the time and effort; this in turn places the onus on the planning department to ensure that the evaluations are intensively utilized.

Comparison with Chile's M&E System

There are other notable issues, particularly if Colombia's M&E system is compared with that of Chile. One is that Chile's system is managed by the budget directorate, ensuring the close integration of M&E and budget work. But in Colombia the M&E system has been essentially managed as a stand-alone activity that almost coincidentally happens to be located within the planning department. Until now the M&E system has not been integrated with either the planning work or the budget responsibilities of the department; nor has there been collaboration with the finance ministry's budget work.

This scenario might be about to change. The rapid evaluations being piloted involve the M&E and budget directorates in the planning department, as well as the budget area of the finance ministry. If these pilots are judged successful—that is, as being cost-effective for purposes of performance budgeting—then it is likely that M&E will become mainstreamed in the core budget work of both ministries.

Some observers have argued that a major constraint on the use of M&E information to influence budget decision making in Colombia arises from the various rigidities in the national budget. These include a large number of earmarked expenditures that cannot easily be varied. As a result, between 90 and 95 percent of budget spending in Colombia cannot be changed in the short term. Various counterarguments can be made here (see annex B); but perhaps the most telling observation is the

Table 7.1: Strengths and Weaknesses of Colombia's M&E System

Strengths	Weaknesses
<ul style="list-style-type: none"> • Very high level of utilization of monitoring subsystem by the president and his office • Performance information used to set performance targets for ministers and their ministries and agencies. Public reporting of the extent to which performance targets are achieved; where they are not achieved, managers have to provide public explanations. • Evaluations conducted externally in a transparent process and considered highly credible by other ministries and the Congress • Evaluations planned and conducted in a collaborative approach involving the planning department and sector ministries and agencies • All M&E information reported publicly and sent to Congress. 	<ul style="list-style-type: none"> • Low-level utilization of M&E information by the budget and planning directorates of the planning ministry, and by the finance ministry • Concerns about the reliability of monitoring data supplied by sector ministries and agencies • Excessive reliance on donor funding of the evaluation agenda

high level of influence of Colombia's *Familias en Acción* evaluation on the government's priorities and on its budget allocations (box 7.2).

One final point when comparing Colombia and Chile is the high level of collaboration between Colombia's M&E directorate in the planning department and the sector ministries and agencies whose programs are being evaluated. The DNP has shared its evaluation expertise and some funding with its sector partners, and none of the sometimes antagonistic relationships that can exist in other countries when the finance ministry plays an active role in the evaluation agenda appear to have arisen. This collaborative approach between the central ministry and the sector entities might change if and when Colombia adopts a more active form of performance budgeting.

However, if performance budgeting is to work well, there will need to be close collaboration among the M&E directorate within the planning department, the budget directorate of that department, and the budget directorate of the finance ministry. One way to foster close collaboration

between these central ministries and with sector entities would be to include all of them in the high-level government committee that has formal responsibility for oversight of SINERGIA. The government plans to move in this direction.

A summary list of the strengths and weaknesses of Colombia's M&E system is presented in table 7.1. The World Bank's diagnosis of Colombia's system is reproduced as annex B to this volume.

Conclusions

The government of Colombia has succeeded in creating a monitoring subsystem of government performance relative to all 320 presidential goals and the country's other development goals. It is notable that the president uses this subsystem intensively in his direct oversight of ministerial and ministry performance and in reporting to civil society.

The government has also embarked on an ambitious evaluation agenda; evaluation findings have already had some significant influence on government decisions and budget allocations.

A weakness in Colombia's system is that M&E information is not yet systematically used for the core budget and planning work of the two ministries responsible for the national budget. There is a good chance that mainstreaming of M&E into budget analysis and decision making will occur, but until it does there will be limits on the extent of utilization of monitoring information and evaluation findings. Such utilization would also considerably increase the probability of sustainability of the central M&E system.

Several ministries and agencies conduct M&E for their own internal purposes, to aid their own planning, analysis, and ongoing program management. These entities are the exception, however. With World Bank support, the government is currently examining options for mandating

M&E more widely within all ministries, agencies, and subnational governments.

It would be fair to conclude that Colombia's M&E system is less well developed and mature than that of Chile. The Chilean system was progressively developed over more than a decade; and although Colombia's system has existed in some form since the early 1990s, it is only since 2002 that it has really blossomed. The Colombian M&E system currently costs around \$2 million per annum, or almost three times the cost of the Chilean system (\$0.75 million per annum).⁴ The source of this cost difference is the particular emphasis on major impact evaluations in the Colombian system. Both countries' systems could be improved, but both—and particularly the Chilean—can be judged to be cost-effective.



8

Australia

Australia elected a reformist government in 1983. That government faced a very difficult macroeconomic situation. In response, it progressively reduced total government spending, from 30 percent of gross domestic product in 1984 to 23 percent in 1989—a very significant reduction by international standards.

At the same time, the government targeted its spending much more tightly toward the most disadvantaged in society. It was keen to obtain greater value for money from government spending, and with this goal in mind it introduced a series of innovative public sector reforms, particularly in the areas of financial management and budgetary reform. Collectively, these reforms placed Australia at the forefront of OECD countries in terms of public sector management.¹

The reforms provided much greater autonomy to the heads of line departments to manage their budget appropriations—under a philosophy of letting the managers manage. Although it had initially been hoped that the reforms would encourage departments to closely manage and measure their performance, this expectation was not met. With the support of other central departments, the powerful Department of Finance (DoF) therefore developed a whole-of-government evaluation strategy. This strategy received the cabinet's strong endorsement through a formal cabinet decision. The strategy followed the philosophy of making the managers manage.

Main Components of the M&E System²

The Australian M&E system largely comprised a formal strategy for evaluations. It was based on a 1988 diagnostic review of evaluation practices in departments and of the overall level of evaluation activity in government. The strategy itself was progressively developed over several years (1987–1991). The strategy had three principal objectives: to encourage program managers to use evaluation to improve their programs' performance; to aid the cabinet's decision making and prioritization, particularly in the annual budget process, when a large number of competing proposals are advocated by individual ministers; and to strengthen accountability in a devolved environment by providing formal evidence of program managers' oversight and management of program resources.³

A centerpiece of the strategy was evaluation planning. This was done through formal portfolio evaluation plans (PEPs), which had to be submitted annually to the minister for finance (see box 8.1). These PEPs had a rolling, three-year coverage and indicated which programs or subprograms would

Box 8.1: Australia's Whole-of-Government M&E System

The main components of the M&E system were as follows:

Formal evaluation planning through PEPs, which listed the government programs the ministry intended to evaluate and the issues to be addressed in each evaluation. These PEPs were prepared annually, on a rolling three-year basis. They included major evaluations only, that is, evaluations of programs considered strategically important to the government: programs with large budgets; those of particular policy importance; problem programs; and pilot programs. The evaluations were conducted by the line ministry itself, usually with some sort of involvement by the finance ministry.

Requirement for every program to be evaluated at least once every three to five years. In practice, this meant that some aspects (such as particular subprograms) of each program were evaluated; most evaluations did not attempt to comprehensively address all aspects of a program's performance. In addition, there was also a requirement that all completed evaluations be published, with the exception of those with national security or industrial relations sensitivity.

Reviews of each ministry's program objectives and performance reporting. These reviews were conducted jointly by each ministry and the finance department, on a rolling basis over a three-year period.

be evaluated, which aspects evaluated, and when. There was a formal requirement that every program be evaluated every 3–5 years. These PEP evaluations were classified as major evaluations. Departments were also expected to initiate other, smaller evaluations, purely for internal management purposes. By the mid-1990s, about 160 PEP evaluations were under way at any one time.

The key issues for the PEPs were the choice of programs to be evaluated and the specific questions each evaluation would address; thus, the ToRs for each evaluation were crucial. These issues were decided through negotiations between the line departments and the DoF's budget sections. For the weaker departments, finance's priorities would largely prevail. For more powerful line departments, the balance of power was more even. Unresolved disputes concerning evaluation priorities would be escalated to the level of ministers or even to the cabinet if agreement could not be reached.

The line departments were responsible not only for evaluation planning, but also for the conduct of evaluations of their programs. Some of the larger departments—such as the employment department—had an evaluation branch with some 20–25 staff responsible for planning evaluations,

providing advice on evaluation methodology, participating in evaluation steering committees, and conducting major evaluations.

Other departments had only small evaluation units, as part of a planning/coordination branch, and devolved the evaluation function to line program areas; they in turn would be responsible to the top management of the department for the quality and rigor of the evaluation (especially for smaller evaluations, which were often in the nature of rapid reviews). A number of evaluations were contracted out to individual consultants or consulting firms.

The DoF's budget sections also involved themselves in the conduct of individual evaluations wherever possible. This would often involve membership of interdepartmental steering committees for the major evaluations; these committees would usually include other relevant sector ministries and other powerful central departments such as the treasury and the department of the prime minister and the cabinet. This broad membership would allow the budget officials (1) to seek to influence the conduct of evaluations, to ensure that problems were fully investigated in an impartial manner, and (2) to comment on draft evaluation reports.

One feature of Australia's M&E system is that a broad definition of evaluation was used. Evaluation was defined as a form of disciplined inquiry: it included rapid evaluations, formal policy reviews, rigorous impact evaluations, and performance audits conducted by the national audit office. The cost of these evaluations varied widely: a sample of evaluations analyzed by the finance department ranged in cost (in 1993 prices) from about \$43,000 to \$430,000.⁴

The evaluation strategy was strengthened in 1991 in response to a performance audit that found that departments varied in their level of commitment to evaluation. The audit office report criticized some departments for poor choice of programs evaluated and for an insufficient focus on the effectiveness of government programs. In response, the DoF created a specialist evaluation unit responsible for providing advice, support, training, and encouragement to other departments, as well as to the budget areas of the DoF itself. This unit also monitored departments' evaluation planning and the number of evaluations conducted; the head of the finance department used this information to informally pressure line departments to improve their evaluation activities.

Australia's M&E system essentially stressed *evaluation*, which was viewed as providing the necessary in-depth, reliable information on the efficiency and effectiveness of government programs. Performance information was also understood to be important, but it was viewed as an issue for line departments to manage. By the mid-1990s, however, the finance department was concerned about the quality of this information and commissioned reviews of departments' annual reports and their budget documentation (which is tabled in the Parliament).

The deficiencies in these reports led the finance department in 1995 to mandate a rolling series of detailed reviews of each department's program objectives and performance information. These reviews were conducted jointly by the finance department and the line department, and recommendations for improvement were required to be implemented.

Performance Budgeting

While Australia's M&E system had three stated objectives, from the perspective of the finance department—which was the primary architect and the overseer of the M&E system—the objective to which it devoted most attention was to support the cabinet's decision making during the budget process. The DoF played a highly influential role in the budget process in Australia. It prepared policy analyses of all new spending proposals prepared by sector ministries, and these analyses accompanied the spending proposals sent to cabinet ministers for their collective consideration when making budget decisions. The DoF thus provided an independent policy analysis that typically constituted a view counter to that of the spending ministry. The work of DoF's budget sections also included the preparation of “savings options”: policy proposals to cut or abolish existing government programs.⁵

The budget process entailed a “marketplace of ideas.” In this inherently adversarial situation, having evaluation findings available about the performance of programs was an important means of ensuring a reliable evidentiary basis for budget decisions. Evaluations had the potential to provide a competitive advantage to those who relied on them. Thus it was important that the DoF's budget sections were fully involved in the evaluation planning of ministries and in the conduct of major evaluations. This ensured that the DoF budget officials were familiar with the quality and any limitations of the evaluations, were fully aware of the evaluation findings and recommendations, and were thus able to use them in their policy analysis work. Involvement of these officials in the evaluations would also increase their knowledge of the program's objectives and the realities of its operating environment—this understanding is important for their work.

It is worth noting here that finance ministries may not always be supporters of reforms designed to strengthen the amount of information available on government performance. Before the reforms, the Australian DoF was heavily involved in the detailed scrutiny of departments' spending activities. The danger is that this traditional focus on

spending can mean that relatively little attention is paid to the results of that spending. And powerful finance ministries can even act as roadblocks to reform. Having the DoF responsible for evaluation oversight ensured that there was a direct influence on the divisional units within DoF that had oversight for line departments.

However, achieving the needed cultural change within the DoF was a slow process over a number of years and involved a substantial staff turnover. The DoF's greater focus on issues of value for money (rather than on spending issues) flowed through to the nature and quality of policy advice the DoF provided to the cabinet; that advice increasingly drew on available evaluation findings.

The most important feature of Australia's M&E system is the significant use made of evaluation findings to support the cabinet's budget decision making. The DoF conducted surveys of its budget staff each year to ascertain the influence of eval-

uations on each new policy proposal of line ministers and the savings options proposed by DoF (or by line ministers if they wished to fund new policies).⁶ By 1994–95, about \$1.75 billion (or 77 percent) of new policy proposals were judged to have been influenced by the findings of an evaluation, and in most cases the influence was judged to be direct. The corresponding figures for savings options were \$380 million (65 percent of the total).

Evaluation findings influenced not only the policy options put forward for the cabinet's consideration, but also the cabinet's decisions. DoF budget officials were also surveyed regarding the extent to which evaluation had influenced the cabinet's decisions in the 1993–94 and 1994–95 budgets. The evidence is mixed, but it indicates that evaluation played a substantive role. In 1994–95, evaluation was judged to have influenced the cabinet's decision in 68 percent of the \$2.846 billion of proposals considered (new policy proposals plus savings options).

Box 8.2: Influential Evaluations in Australia

In the 1996–97 budget the new government was determined to both reduce and reprioritize government spending. Particular focus was given to labor market and related programs, which accounted for \$2.90 billion in spending annually. The minister for employment articulated the government's overall policy goal as being to provide assistance to the long-term unemployed and to those at risk of entering long-term unemployment. This focus was adopted both for equity and efficiency objectives, such as achieving a better match of labor supply and demand. At the same time, the minister wanted to achieve better value for money from labor market programs in the tight budgetary environment.

Australian and international evaluation findings were drawn on heavily to help guide the policy choices made. The minister highlighted the relative cost-effectiveness of different labor market programs. A key measure of this was estimated by calculating the net cost to government for each *additional* job placement from different programs—as measured by the *increased* probability of an assisted person being in a job some 6 months after he or she had participated in a labor market program. (The base-

line was a matched comparison group of individuals who did not participate in a program.)

Evaluation findings showed that the JobStart program, which provided wage subsidies, had a net cost of \$3,700 per additional job placement, whereas the JobSkills program, which was a direct job-creation program, had a corresponding net cost of \$57,800. The minister noted, "The government will be . . . concentrating its efforts on those programs which have proven most cost-effective in securing real job outcomes." As a result, the JobStart program was retained and the JobSkills program was substantially scaled back and more tightly targeted to job seekers who were particularly disadvantaged.

Total savings to the government from its reduction and reprioritization of labor market programs were about \$1.14 billion over two years. The cabinet also commissioned a series of major evaluations of its new labor market programs and of the new arrangements for full competition between public and private employment service providers.

Source: Mackay 1998a.

Table 8.1: Strengths and Weaknesses of Australia's M&E System

Strengths	Weaknesses
<ul style="list-style-type: none"> • Evaluation findings heavily used for budget analysis, policy advice, and by the cabinet in its budget decision making • High utilization of evaluation findings by sector departments and agencies • Evaluation conducted as a collaborative endeavor between finance department, other central departments, and sector departments 	<ul style="list-style-type: none"> • Uneven quality of evaluations • Insufficient availability of advanced evaluation training • Insufficient attention to regular performance information • A claimed administrative burden on departments

Source: Mackay 1998a, 2004.

The corresponding proportion for the 1993–94 budget, however, was only 19 percent of proposals. One important reason for this difference was the substantial revision of labor market, industry, regional, and aboriginal policies in the 1994–95 budget—the major policy review on which these decisions were based had been heavily influenced by a number of evaluations commissioned specifically to help guide the policy review.

Only a few programs were terminated as a result of an evaluation; given the emphasis in the budget process on portfolio (defined as a line department plus outrider agencies) budget envelopes—in effect, portfolio budgeting—any program termination would often result in spending reallocation within that portfolio. There were major instances where programs were significantly cut as part of a major reprioritization, for example, in the labor market and social security areas; these cuts reflected the desire to maximize value for money given the government's policy priorities (box 8.2).

Strengths and Weaknesses of Australia's M&E System

Strengths

Some commentators have observed that “program evaluation (in Australia) has been applied more extensively and systematically than in any other country” (Kim and others 2006). The outstanding feature of Australia's M&E system was the

high utilization of evaluation findings in the budget process, as a key input for both high-quality policy advice (Uhr and Mackay 1996) and for the cabinet's budget decision making (table 8.1). As noted in the chapter 5 discussion on performance budgeting, the influence of evaluation findings on the government's decisions was typically indirect—ranging from a major influence in a number of cases to little or none in others. This is essentially the most that can be expected from a whole-of-government M&E system, in terms of potential utilization of the M&E information that the system produces.

The Australian auditor-general has commented, “In my view, the success of evaluation at the federal level of government . . . was largely due to its full integration into the budget processes. Where there was a resource commitment, some form of evaluation was necessary to provide justification for virtually all budget bids” (Barrett 2001).

Another strength of the M&E system was the high level of utilization of the information by line departments and agencies. A performance audit conducted by the Australian National Audit Office (ANAO) in 1997 concluded that line departments were making considerable use of their evaluation findings to help improve operational efficiency. To a lesser extent, they were also using these findings to help guide their own resource allocation decisions and in the design of service

quality improvements (Australian National Audit Office 1997).

This high level of utilization by line departments reflected another strength of the Australian M&E system: evaluation was essentially a collaborative effort involving the DoF, other central departments, and the line departments. Although responsibility for evaluation was largely devolved to line departments, the involvement of the central departments in the planning and oversight of the major PEP evaluations helped achieve broad ownership of the evaluations themselves and of their findings.

This approach stands in contrast to that of Chile, for example, where the line ministries generally have little or no ownership of the evaluations the finance ministry commissions. The Australian approach also ensured that line departments' deep knowledge and understanding of their own programs were used intensively in evaluations; a danger with an evaluation system that relies on externally conducted, independent evaluations is that it can fail to draw on this program expertise.

Weaknesses

The downside of Australia's more collaborative approach, however, was an uneven quality of evaluations conducted by line departments. The ANAO performance audit analyzed a sample of evaluation reports and concluded that more than one-third suffered from methodological weaknesses of one kind or another.

One reason for this was that many of the program areas of line departments that had responsibility for conducting or commissioning evaluations lacked sufficient skills to ensure that high-quality evaluations were conducted. The DoF provided basic training in evaluation methods and issued handbooks on program evaluation and cost-benefit analysis. But the ANAO audit reported that 20 percent of line departments were concerned about the lack of available training in advanced evaluation techniques.

In retrospect, one option to address this issue of evaluation quality would have been for the DoF

to mandate creating sizeable central evaluation units in each department. Another option would have been to centralize the entire evaluation function in the DoF—this would have required the creation of a very large evaluation office and would have been contrary to the devolutionary nature of most of the public sector reforms.

Two other weaknesses of the Australian M&E system are worth noting. First is the relatively weak emphasis given to the regular collection and use of performance information. The DoF explicitly advised departments of the importance of developing sound program objectives and having sound performance information—not least to facilitate the conduct of evaluations. It was not until 1993, some six years after the evaluation strategy was initiated, that the DoF commissioned broad reviews of the quality of departments' annual reports and budget documentation.

Following these reviews, which were critical of the quality of program objectives and performance information, the DoF mandated a series of in-depth, rolling reviews of all departments' and agencies' program objectives and performance information. It required that action plans be implemented to address any problems identified. One lesson from this experience is that all M&E systems require ongoing review and adjustment; it is necessary to monitor and evaluate an M&E system, just as it is necessary to monitor and evaluate any other type of public sector reform.

A final weakness of the M&E system claimed by some departmental secretaries was that the formal requirements for evaluation planning and reporting were too burdensome. Some departments did create complex internal processes for evaluation planning and ended up preparing 120-page PEPs; yet the DoF guidelines did not mandate such complexity, and the guidelines recommended much smaller PEPs as good practice.

The Evaluation System—A Postscript⁷

The election of a conservative government in 1996 led to the considerable downsizing of the civil service; a substantial weakening of the policy advising process and the role of the DoF, especially

in the budget cycle; and the dismantling of many of the remaining central controls and requirements. At the same time, a much higher level of autonomy was given to the heads of line departments—a return to the philosophy of letting managers manage.

Although line departments are still required to report their performance publicly to the Parliament, the DoF now applies no central controls or quality standards for M&E. As part of these changes, the decade-long formal evaluation strategy—and thus the government’s M&E system—was dismantled. The Australian auditor-general has characterized these reforms as a deregulation of evaluation. While there remain some line departments that can be considered to be good practice “islands” of M&E,⁸ in terms of their conduct, quality, and use of M&E, these appear to

be the exception rather than the rule.⁹ ANAO performance audits have highlighted the poor quality of performance information that most departments now provide to the Parliament.

Although the DoF still provides advice on departments’ budgets, it lacks systematic, reliable monitoring information and evaluation findings on which to base this advice. As the OECD has concluded—

In Australia, the deregulation of the public service and the adoption of an arm’s-length posture by the central agencies allowed management freedom but is currently considered to have deprived the Finance Ministry of the information necessary for it to adequately advise the Minister (OECD 2002).



The Special Case of Africa

The experience of African countries is relevant to poor countries in other Regions, especially those preparing poverty-reduction strategies. Africa also provides lessons on how to build M&E capacities incrementally, especially when there is the possibility of intensive donor assistance. These lessons are also relevant to middle-income countries, such as those that are not yet committed to comprehensive improvement of their M&E systems. In such countries, a more cautious focus on incremental changes can be appropriate if there is the potential to demonstrate that M&E is a cost-effective government activity.

It is widely accepted that the extreme poverty situation facing most African countries provides a clear priority for intensive development support. More than 30 African countries have prepared an interim or final PRSP, a document required for access to debt relief under the Heavily Indebted Poor Countries Initiative.¹ PRSPs set development targets and are intended to report on results achieved. In practice, this has meant a focus on the extent to which a country has achieved the MDGs.

This puts a premium on having adequate national statistics, which in turn is leading to intensive donor support for statistical capacity building, such as assistance for population censuses and household surveys. Particularly with their national statistical offices, countries appear keen to accept this support. PRSPs usually present their national monitoring (that is, statistical) systems as

synonymous with M&E, and the need to give priority to M&E has become a mantra that is widely accepted by governments and donors alike. In many cases, however, national monitoring systems are principally designed to meet donor data requirements (IEG 2004b).

Moreover, PRSPs end up focusing on the amount of budget and other resources spent on national priorities and national progress against the MDGs. These two issues are certainly important, but what is absent from this focus is what Booth and Lucas (2001a, 2001b) have termed the “missing middle”: performance information on the intervening steps in the results chain, involving government activities, outputs and services provided, and their outcomes; and in-depth evaluative evidence linking government actions to actual results in the field.

Statistics on amounts spent and on poverty levels are both very important, but unfortunately neither is able to measure the government's performance in terms of the *results* of its spending—the outputs, outcomes, and impacts of the government itself. MDGs and other measures of poverty provide a bottom-line measure of *country* performance but fail to reveal the contributions of the government compared with donors, the private sector, and civil society groups such as NGOs.

Most African countries are simply too poor to be able to conduct evaluations and reviews; they rely instead on donors for such work. A difficulty is the heavy burden placed on countries to meet the M&E requirements of donors, in terms of inspection missions, provision of performance information, unharmonized donor evaluation criteria and methods, and so forth (IEG 2003a). Lack of donor harmonization has imposed a heavy burden of information supply on aid-dependent countries. However, donor cooperation and harmonization can be facilitated through sectorwide approaches. In Tanzania, for example, there is a health sector working group, made up of government and donors, that not only analyzes sector performance and policies but has also reviewed sector M&E systems and identified M&E capacity-building priorities. This sector working group also commissions evaluation or research into selected issues (Stout 2001).

The move toward greater use of programmatic lending to countries provides another way to mitigate the harmonization problem, because it reduces the scope for project-specific M&E and thus the scope for balkanized donor M&E. In Uganda, for example, the World Bank and other donors provide programmatic budget support to the government. Such programmatic support is becoming increasingly common in African and other debt-relief countries.

Lessons from Uganda²

Some African governments, such as Uganda and Tanzania, understand well the importance of having reliable and comprehensive performance

information available. They use this information intensively in preparing their national plans and in determining budget priorities (see, for example, Government of Tanzania 2001; Ssentongo 2004; Government of Uganda 2004, 2006). A notable feature of both countries is that their national plans—in Uganda they are known as Poverty Eradication Action Plans (PEAPs)—predate the PRSP initiative. This experience made it easy for both countries to prepare PRSPs; indeed, they simply had to prepare summarized versions of their national plans to meet the PRSP requirement.

Uganda has had a number of M&E initiatives and systems. It was, for example, the first country in which PETS were undertaken (box 9.1). However, diagnoses of Uganda's M&E arrangements in 2001 (see table 9.1) and 2003 revealed a large number of uncoordinated and unharmonized monitoring systems at the sector and subsector levels—at least 16 separate systems (Hauge 2003). In addition, a detailed investigation of three sectors (health, education, and water and sanitation) revealed a considerable data-collection burden at the district and facility levels.

The management information systems for those three sectors collected data on nearly 1,000 performance indicators, involving almost 300,000 data entries per annum for each of the 110 districts in Uganda. These indicators largely focused on spending, activities, and the physical state of facilities such as schools and health clinics.

However, measures of client satisfaction and outcome measures, such as health status and learning outcomes, were largely missing. Unfortunately, the quality of the data was highly uncertain and often considered poor. As a result, the sector ministries and agencies relied heavily on inspection visits rather than on self-reported performance indicators. Hauge and others (2002) estimated that site inspections in the health sector alone were costing the equivalent of 1,400 staff years per annum, often consuming the time of qualified medical personnel.

The diagnostic findings concerning the multiplicity of M&E systems and performance indica-

Box 9.1: An Influential Evaluation in Uganda

In the early 1990s, Uganda, like many other developing countries, was concerned with the poor performance of public services such as education and health. A major cause was believed to be the “leakage” of funds that did not reach front-line agencies. With World Bank support, the PETS methodology was developed and applied in the primary education sector in 1996. Its purpose was to measure the proportion of funds provided by the central government that reached primary schools and the extent of the leakages and to recommend ways to reduce them.

The PETS analyzed the timing of budget flows through various tiers of government and compared budget allocations to actual spending on primary schools. As adequate public accounts were not available regarding actual spending, surveys were conducted in 250 government primary schools in 19 districts, and a panel dataset was created on spending and outputs for 1991–95. The study found that only 13 percent of earmarked (nonwage) funds actually reached schools in 1991–95. The remaining 87 percent disappeared or was used by district officials for other purposes. About 20 percent of funds allocated for teacher salaries went to “ghost workers” who did not exist or who were not working as teachers.

The study findings attracted considerable media attention, and the government decided that the information on the amount of funds

allocated to, and received by, each school should be widely disseminated through local newspapers and radio stations and publicly displayed at each school. This helped ensure that parents became aware of the funding situation facing their child’s school. This provided the information they needed to hold teachers and school principals accountable for lack of available teachers or for an inadequate supply of textbooks. These steps also demonstrated to local governments that the central government had resumed its oversight function.

Two follow-up PETS showed that the flow of nonwage funds reaching primary schools had improved from 13 percent in 1991–95 to between 80 and 90 percent in 1999–2000. This example demonstrates that quantitative data on public services are a powerful tool for mobilizing civil society’s “voice.” Although individual complaints can be brushed aside, public feedback backed by systematic comparative data is difficult to ignore and can provide a spark for public action.

The first PETS cost \$60,000 and has been estimated to have helped increase the amount of funds reaching primary schools by more than \$18.5 million per annum. This indicates that PETS are a highly cost-effective evaluation tool. The government of Uganda now routinely conducts PETS for each basic service sector.

Source: Bamberger, Mackay, and Ooi 2004, 2005.

tors—and the heavy burden imposed on frontline staff—caused considerable consternation within the government and led to the decision to create a National Integrated M&E System (NIMES) under the aegis of the Office of the Prime Minister. The objective of NIMES is to create an umbrella M&E system within which existing systems will be coordinated and harmonized and government capacities to conduct and use M&E strengthened (Government of Uganda 2004, 2006).

Various working groups have been created under NIMES addressing the following issues: M&E in local governments; policy research; evaluation; national statistical data; sector management information systems and spatial data; civil society organizations and M&E; and financial information. At least four donors provide funding support for NIMES, in the amount of some \$7.4 million

over a three-year period, in addition to the government’s own funding.

NIMES is reducing the very large number of performance indicators, especially at the sector level, with a greater focus on outputs, outcomes, and impacts, as well as on the setting of targets. The World Bank and seven other donors now prepare a joint strategy for providing support to Uganda,³ based on the national plan (PEAP) goals and objectives, and they rely heavily on the government’s own monitoring information to assess their own performance. Donors still conduct their own separate evaluations, however.

In addition to the NIMES, the government has also embarked on a process to improve its performance and accountability. This includes the development of results-oriented management in public

Table 9.1: Uganda M&E—Summary Diagnosis of Strategic Issues, Challenges, and Possible Actions

Monitor and provide feedback on progress in poverty alleviation							
Strategic M&E issues	1. Coordination and harmonization	2. Development goals, targets, and performance indicators	3. Incentives to contribute to results	4. Devolution of managerial autonomy	5. Role of civil society in enforcing transparency and accountability	6. PEAP partnerships principles	7. M&E skills training
Positive elements in Uganda	<ul style="list-style-type: none"> • Draft poverty monitoring strategy • Sector working groups as nexus of planning, budgeting • Efforts to harmonize project progress reporting • Earmarking of five percent of poverty action funds for monitoring and accountability 	<ul style="list-style-type: none"> • PEAP/PRSP as overall framework of poverty priorities • Training in results-oriented management being piloted • “Indicator retreat” as part of budget cycle 	<ul style="list-style-type: none"> • Recognition of service delivery effectiveness as imperative of public management • NSDS 2000 	<ul style="list-style-type: none"> • Decentralized responsibility for service delivery • Introduction of output-oriented budgeting • Comprehensive district plans • Local government development plan capacity development 	<ul style="list-style-type: none"> • Consultative nature of PEAP process • Transparency of budget process • Practice of public notices • Significant capacity of NGOs • Government–civil society dialogue at central level 	<ul style="list-style-type: none"> • Draft CDF partnership principles • Trend toward budget support • Consultative group meeting scheduled as part of budget cycle 	<ul style="list-style-type: none"> • Awareness of M&E importance • Availability of local researchers and local academic and training traditions
Some challenges	<ul style="list-style-type: none"> • Separate planning and reporting formats for different funding sources • Sector/district policies, budgets, work plans approached as separate exercises • Alignment and coordination of different results management initiatives • One-third of official development technical assistance is outside of government budget 	<ul style="list-style-type: none"> • Inconsistent clarity of goals at sector level • PEAP goals correspond to ministerial activities rather than to poverty outcomes • Few goals are defined with measurable timeframe, baseline, or targets • Weak linkage between sector and district goals 	<ul style="list-style-type: none"> • Performance assessed in terms of expenditure and bureaucratic activity • Weak linkage between resource allocation and performance • Rewards geared to good paperwork • Inconsistent enforcement of sanctions for poor performance • Corruption largely unpunished 	<ul style="list-style-type: none"> • Generally weak management capacities at local levels • Prescribed spending ratios of conditional grant scheme gives little flexibility for managers to adapt to local needs • Number and level of posts directed from the center 	<ul style="list-style-type: none"> • Need for improved stakeholder consultation in priority setting • One-third of official development assistance is technical assistance outside of government budgets and M&E practices • GoU/NGO dialogue at center is not mirrored at local levels 	<ul style="list-style-type: none"> • One-third of official development assistance is technical assistance outside of government budget and M&E • Nearly 300 stand-alone projects remain • 20 separate annual program reviews • Local donors cannot depart from corporate M&E guidelines 	<ul style="list-style-type: none"> • Weak management skills at local government levels • Likely increase in demand for management and conduct of M&E dealing with interrelationship between service delivery and poverty outcomes

<p>Possible actions to address challenges</p>	<ul style="list-style-type: none"> • Identification of an M&E champion ministry or agency • Establishment of core M&E arrangements (such as through a formal M&E framework), harmonization of terminology, reporting formats, and periodicity • Improved coordination between inspection and audit agencies 	<ul style="list-style-type: none"> • Cascading of PEAP goals and targets through planning, budgeting, and work planning at sector, district, and facility levels • Long-term expenditure framework focus on defining medium-term PEAP goals and targets 	<ul style="list-style-type: none"> • Introduction of reach and outcomes as yardsticks of success and performance reward • Use of the NSDS as barometer of client satisfaction improvements • Introduce value for money concerns in finance act • Introduction of client service charters 	<ul style="list-style-type: none"> • Allow greater local autonomy over recruitment, salaries, and non-wage expenditures • Ensure stronger local oversight as the quid pro quo • Introduce participatory M&E practices as key management function 	<ul style="list-style-type: none"> • Extend transparency practice from allocation to execution • Client report cards as complement to the NSDS • Make NGOs eligible for poverty action fund monitoring and accountability funding • Introduce client service charters 	<ul style="list-style-type: none"> • Leverage of donor support for the CDF and PEAP to increase synergy in planning, reporting, and review • Use poverty-reduction support operation policy matrix as joint planning and review mechanism 	<ul style="list-style-type: none"> • Strengthen local capacity for evaluation skills training, for example, by training trainers at national institutions • Coordinate use of funding earmarked for M&E under poverty action fund, the local government development program, and Economic and Financial Management Programme II • Establish national evaluation association
--	--	---	--	---	---	---	--

Source: Hauge 2001.

Note: CDF = comprehensive development framework; GoU = government of Uganda; NGO = nongovernmental organization; NSDS = National Service Delivery Survey; PEAP = poverty eradication action plan; PRSP = Poverty Reduction Strategy Paper.

service organizations and its links with the budget process and with the staff performance appraisal process. The government's efforts include actions to strengthen both top-level political and civil service commitment and the demand for a greater performance orientation.⁴

Some Conclusions

Uganda and other African countries already possess M&E systems. They also receive donor support for statistical capacity building and have access to donor evaluations. The challenge these countries face is not developing new systems, but rationalizing and improving what already exists. There are problems with data quality and unharmonized donor requirements for M&E—a situation of *too much data, not enough information*. Compounding these problems on the supply side is the fact that in most other countries there is weak government demand for M&E information.

Although it would be unrealistic to expect most African countries to build comprehensive, reliable M&E systems, there are a number of important elements that they could feasibly undertake. What follows is a list from which African and other countries preparing a PRSP could draw, either with or without donor support. The advantages, costs, and limitations of some of these M&E tools

are discussed in *Monitoring and Evaluation: Some Tools, Methods and Approaches* (IEG 2004b):

- Financial management information systems to support better financial tracking of government spending
- Public expenditure tracking surveys to identify “leakage” and to trace the effects of corruption
- Service delivery surveys of client satisfaction and perceptions of the quality of government services
- Rapid appraisals—for example, of “problem” projects or programs
- National and sector statistical collections—especially relating to national priorities such as the MDGs
- Sector ministries’ administrative data.

The only caveat with this list is that, in some senses, less is more. One danger to avoid is the tendency to overengineer whatever M&E system is being created. It is therefore important to carefully monitor the extent to which each type of M&E information is being used. Where utilization is low, it is necessary to identify the reasons, such as low awareness of its existence, a low level of demand for it, poor quality data that are considered unreliable, or a lack of staff able to analyze and act on the information. This helps identify the steps necessary to improve supply or to increase the demand for M&E information.