Monitoring and evaluation (M&E) are means to multiple ends. Measuring government activities, constructing and tracking performance indicators across sectors and over time, evaluating programs—these activities can be carried out and tied together with different objectives in mind. It would certainly be possible to use M&E purely as a way to improve transparency and accountability, by making more information on the workings and results of government programs available to the public. One can also focus M&E on managerial purposes, to reward performance inside ministries and agencies. But surely a crucial element of running an effective public sector would be missing if M&E were not used to inform the spending of public money. This briefing note will introduce the main issues surrounding M&E as a tool for budgeting—a system usually referred to as performance budgeting—to help policy makers make strategic decisions about their M&E systems by outlining different design choices and their respective advantages and pitfalls.

Budgeting is said to have three aims: (1) fiscal discipline, (2) allocative efficiency, and (3) operational efficiency. The most basic question then is: what can M&E do to help officials involved in the budget process achieve those aims? A traditional ministry of finance usually focuses on enforcing fiscal discipline by tightly controlling spending at the line item level. The German budget, for instance, has around 6,000 line items. This level of detail gives the finance ministry many levers to keep ministerial spending in check. However, it has often been criticized as relatively inflexible and inefficient for budgetary decision making. Budget officials have to spend significant time reviewing small items that have little bearing on the important policy decisions, and the system does not facilitate considerations of the results of public spending.

A well-designed M&E system has the potential to overcome some of the shortcomings of traditional budgeting. By defining indicators for public spending outputs and outcomes, governments gain the ability to monitor nonfinancial performance. By carrying out evaluations that turn performance information into implementable policy recommendations, governments can more efficiently allocate resources to programs with the greatest contribution to its policy priorities. In focusing the effort of public managers on performance, M&E is also expected to contribute to better operational efficiency.

Many Organisation for Economic Co-operation and Development (OECD) countries have introduced some elements of monitoring and evaluation into their budget process over the last three decades. These performance budgeting reforms were often part of other efforts to improve performance in the public sector and to make the budget process more flexible for difficult fiscal climates. Early adopters, such as Australia, New Zealand, and the United Kingdom, pursued their reforms most vigorously in times of fiscal crisis. The main motivation was often not primarily to strengthen spending discipline, but rather to develop the tools that would allow the government to make better use of existing funds.

Many middle-income countries have since taken up the ideas first seen in the OECD and adapted them to their contexts. In many lower-income countries, the context for introducing M&E into the budget process is quite different. The fight against poverty and the joint push by
governments and donors to make progress on the Millennium Development Goals provide a results orientation by default. For lower-income countries, developing a link between the poverty reduction goals in the Poverty Reduction Strategy Paper (PRSP) and the annual budget cycle is a key challenge.

What Are the Different Tools, and How Do They Fit Existing Contexts?

It is very hard to imagine a budget process that does not monitor: even the most basic financial management system tracks the flows of resources. A detailed line item budget is an excellent tool to monitor inputs of money. When using the term M&E, however, “monitoring” means the regular measuring of the outputs of government programs and possibly their outcomes as well. If M&E together are used for budgetary decisions, a form of performance budgeting emerges (Arizti et al. 2009; Curristine 2007; Shah 2007).

The most rigid form of M&E is directly linking funding to performance, most commonly outputs. The ministry of finance essentially buys a certain level of performance from spending ministries and only funds what it buys. Formulaic performance budgeting is quite rare and can only be found in a few countries in certain sectors, such as health or education, where outputs can be rigidly standardized. No country has institutionalized such a system for its entire government, because it is both technically and politically almost impossible to automatically tie funding to predefined levels of performance. There is also not much of a role for evaluations in such a system, which is problematic.

Performance information cannot be used to automatically adjust spending, it needs to be interpreted. One might expect programs with negative evaluation results to receive funding cuts, however, politically important programs with negative results might just as well receive funding increases, because failure would be unacceptable and fixing failed programs costs money. Formulaic performance budgeting decreases the scope for informed decision making by public officials.

On the other end of the spectrum, the M&E-budget link can be purely presentational. Sector ministries or the ministry of finance can publish performance data or submit data to the legislature as an appendix to the budget or in the form of performance reports during the budget year. These reports and figures might be used in budget negotiations by legislative or executive participants, but there is no requirement to formally consider them.

Presentational performance budgeting can still raise important issues and make budgetary decision makers aware of problems that a traditional line item budget would not uncover. It can also be the result of a budget system where either the finance ministry or the legislature enjoys a large degree of discretion over how they reach their allocative decisions. In many cases, it may be a transitional state, where proper decision-making structures have not yet been set up, or an unintended system where information is produced somewhere in the system, but not used by key budgetary actors.

In the OECD, countries that adopt performance budgeting tend to develop a performance-informed budget process, where M&E instruments serves as analytical tools. In this model, performance indicators are integrated into the budget and regular evaluations are scheduled into the budget cycle. Performance indicators and evaluations are institutionally considered during budget formulation. This process can take various forms, in the United Kingdom, for instance, the finance ministry uses three-year spending reviews to negotiate spending plans with ministries. The basis of planned expenditures is the “public service agreement,” which uses performance indicators to define targets the ministry is supposed to achieve with a given amount of funding over three years. These targets form the core of the interaction between the finance ministry and spending ministries and are meant to capture what a ministry does and to set the agenda for its future achievements. The 1998 spending reviews contained a total of about 600 headline targets; this number was reduced to 110 by the 2004 budget exercise. Inside each ministry, these targets are broken down along results chains into lower-level targets for operational purposes. Ministries have a very high degree of discretion over the internal allocation of funds to achieve their targets. The result is a highly structured, top-down system where performance indicators are very prominently used to inform budgetary decisions.

Many other countries have developed less ambitious ways to use performance indicators, such as by complementing the traditional line items
in the budget document with data on ministerial performance.

Performance indicators are available as a source of information that can be used for budgetary decisions without laws or regulations prescribing exactly how legislators or budget officials might use them. A singular focus on monitoring performance indicators can be criticized for several reasons. First of all, if indicators are the main sources of information considered for crucial decisions, then incentives grow for public officials to focus on delivering indicators but not necessarily improving the underlying performance. There are numerous examples of public officials trying to “game” indicators to secure funding. Secondly, the supply of performance data can easily outstrip demand. When ministries expect that their performance indicators will shape future funding, they have every reason to develop as many indicators as possible to ensure that every aspect of their work is captured. The result is that much information is produced, but not necessarily used.

In addition, performance indicators alone are not directly actionable. A performance increase of 2 percent can be a substantial success or a deep disappointment when compared to other countries. Numbers can be affected by external factors and changes in funding can influence outcomes with a lag of several years, or very disproportionately, or not at all. It often takes time and specialized policy knowledge to interpret performance information, which are in short supply during the annual budget cycle.

The need to have actionable information is why the combination of monitoring and evaluation is crucial. An evaluation is a systematic review of a program or project that identifies strengths and weaknesses and recommends actions that officials can use for budget formulation. Many policy analysts first think of an evaluation as a rigorous impact analysis that is able to identify the specific contribution of a policy intervention by comparing beneficiaries and nonbeneficiaries. Such impact evaluations are indeed used in budgetary decisions and can have a huge impact, particularly for the funding of new policies. The reputation of conditional cash transfer programs, for instance, was in no small part established by the impact evaluations of Mexico’s Progresa program in the 1990s. These evaluation results allowed public officials to judge the merit of this innovative policy intervention against other programs with similar aims. However, impact evaluations are not appropriate for all budgetary evaluation needs. Most importantly, they are expensive and take time, usually more than one year.

For a frequent and comprehensive evaluation system, quicker, desk-based evaluations are an important tool. Rather than carrying out original research, a rapid evaluation would rely on the available information that can be provided inexpensively from the evaluated program and other public sources. This can cut costs substantially and reduce research time to a few months—a crucial difference, because they can be made to fit into a single budget cycle. A rapid evaluation can be initiated during budget deliberations, commissioned and carried out the same year, with the results available to feed into the budget preparation the following year.

For instance, Chile conducted more than 250 rapid evaluations between 1997 and 2010. The evaluation schedule is decided in negotiations between the finance ministry and Congress. Evaluation teams are chosen from a pool of expert consultants. All evaluation recommendations are monitored by the Chilean budget office and revisited annually during budget negotiations. While not comparable to impact evaluations in depth and academic rigor, desk-based evaluations focus on consistency and operational issues and can uncover significant issues relevant to budgetary decisions.

Countries such as the United Kingdom also use comprehensive spending reviews to consider all available M&E data to evaluate spending in one policy area. The UK Treasury does not just consider its own M&E data, but also evaluation results commissioned by other government bodies, such as the National Audit Office. Spending reviews are carried out every two to three years.

A traditional line item budget is often criticized for its inefficiency and inability to respond to new policy challenges. Budget officials in the finance ministry as well as legislators in parliament have limited time to consider the annual budget. A budget consisting of thousands of individual items makes it possible to closely control where money is being spent. Change, however, is incremental and often limited to distributing the marginal increase from year to year. A combination of soundly measured and regularly monitored performance indicators and accessible evaluations can, in theory, enable budget decision makers to move beyond the limits of incremental budgeting and respond to new policy challenges.
to larger fiscal challenges. But “in theory” is a big caveat, because in practice, M&E only makes a difference if the demand is right.

The Institutional Roles—Who Does What?

M&E is about supply and demand. Unless actors in the budget process demand and use the outputs of M&E, even the most sophisticated tools will have no effect. It is also of critical importance that budget processes are inherently political in nature. Each phase of the budget process has its technical underpinnings, but ultimately it is about turning political priorities into administrative reality. A country that successfully introduces M&E into the budget process does so because enough actors demand the information that M&E provides.

The finance ministry is the main agency in charge of the budget process and often the appropriate ministry in which to anchor the M&E system. Chile is a good example; the Chilean budget process is possibly the most centralized among OECD countries. The budget director is a key advisor to the president, and the budget process is tightly managed by the budget office (DIPRES). The natural result of such a budget system is that the M&E system is also managed by the budget office. The technical mandate of DIPRES is solidly defined in a legal framework that gives it considerable discretionary scope to change budgeting procedures without new legislation. The budget office controls the schedule, the methodologies, ensures quality, and manages the flow of information between agencies, ministries, itself, and Congress.

Yet even in Chile, the budget office is mostly a compiler of, not producer of, information. Evaluations are commissioned by the government, but their implementation relies on a network of universities and other researchers. The indicators that feed into the monitoring system are supplied by agencies and ministries themselves. The main role of the budget office is quality control. It bears mentioning that the budget office strongly supports the use of M&E information both throughout the executive bureaucracy and for legislative purposes. Yet international studies point out that the budget office remains the principal user and supporter of M&E in Chile. This system has worked very well in Chile over the past 15 years, but it is not the only institutional setup that works. Very few other countries can rely on such a strongly centralized budget process to underpin the introduction of M&E.

The main design challenge is how to tailor the system to the likely users of M&E information and their position within the budgeting framework. This is a universal concern, irrespective of a country’s stage of development (see box 1). There is a growing body of literature on fiscal institutions that points out that the centralization of the budget process depends on macropolitical factors. The most important factors relate to electoral and party systems and the constitutional separation of powers between the legislature and the executive branch (Hallerberg 2004). They are thus largely beyond the control of technical policy makers. In countries where the right preconditions do not exist, other successful models can be viable alternatives to centralization, but only when they take the institutional balance into account. In other words, there is little reason to expect that a highly centralized M&E system will produce good results if the budget system itself is not centralized.

Centralization is most likely in parliamentary systems where either one or a few closely aligned parties form the government, such as in the United Kingdom. In the United Kingdom, the party

---

Box 1: Mind the Gaps between the Poverty Reduction Strategy and the Budget

A very different challenge arises in the many countries that tried to implement poverty reduction strategies (PRSs) to align country and donor efforts toward reaching the Millennium Development Goals. PRSs are often drafted by planning ministries or special commissions without a central role in the regular budget process. Critics noted that these strategies were geared toward donor audiences and not properly implemented. The format and planning schedule of PRSs did not necessarily fit well into the existing budget process. Countries like Uganda, which was quite successful at using the PRS for planning and budgeting, did so by putting the budget office firmly in charge. The M&E systems that Uganda implemented to feed into the PRS were designed with the budget process in mind.

with a majority in parliament automatically heads the executive as well, which makes for relatively clear lines of accountability within government. If the government has resolved to employ M&E in the budget process, and already delegates strong budgetary authority to the finance ministry, then the budget office becomes the main consumer of M&E information. When legislators discuss the draft budget in parliament, their majority will be part of the governing coalition and their policy priorities should be broadly aligned with those of the finance minister.

Many finance ministries make M&E information available to legislators to build legislative support for the M&E system in general and to support controversial budgetary decisions that might be based on M&E data. Legislators have an incentive to consider M&E information, especially when budgetary decisions affect their own constituency. In fact, in many parliamentary systems, the supreme audit office is closely tied to the legislature and often carries out its own evaluations independently of the executive branch. Nevertheless, the budgetary relationship between the budget committee in parliament and the finance ministry is not adversarial.

In presidential countries, the constitutional balance between the legislature and the executive branch plays a crucial role. Chile is a fairly exceptional case in this regard, because the powers of Congress to amend the president’s budget proposal are severely limited. On the other end of the continuum lies the United States, where the president’s budget can be famously “dead on arrival” if the Congress is controlled by the opposing party. Most presidential systems lie somewhere in between (see box 2).

Irrespective of the constitutional or legal powers of the finance ministry, many budget offices find it useful to give agencies and ministries a substantial role in the M&E system. In some countries, a tradition of collegiate cabinet government makes it difficult for the finance ministry to assume tasks without being seen as encroaching on the authority of other ministers to shape the policy in their respective sectors. To push ahead without properly involving sector ministries can imperil the sustainability of reforms, a problem even for successful reformers like Chile.

Generally, finance ministries make their own workload more manageable by delegating parts of M&E to ministries. If ministries operate performance indicators and buy into evaluation results, they can forestall many potential conflicts during budget negotiations with the finance ministry and shape their budget proposals in a way more acceptable to both sides. A strong stake for ministries and agencies also makes it more likely that M&E results are used for operational and managerial purposes within ministries, something that finance ministries might reasonably expect to benefit operational efficiency.

When Australia introduced its evaluation system in 1987, ministries were obliged to develop

---

**Box 2: Not All Presidential Countries Are the Same: The Experience of Mexico**

In Mexico, congressional budget powers are considerable. Since 1998, three parties of roughly equal size have shared the legislature. By default, the president’s party does not have a legislative majority and thus cannot pass the budget alone. When Mexico moved toward developing M&E capacity in the public sector, it faced a potential impasse. On the one hand, the legislative power of the purse made Congress a legitimate and inevitable consumer of M&E data. The legal framework of budgeting also requires congressional approval for the necessary legal changes to the budget process. Yet if the M&E system was led by the legislature, there would be a strong incentive for the executive bureaucracy not to develop a stake in M&E for fear that the new tools would expose them politically. On the other hand, the Ministry of Finance could not fully centralize M&E authority either, because to do so might jeopardize the credibility of M&E information in Congress. This issue was resolved by vesting much technical expertise and authority over performance monitoring and the evaluation program in an independent technical body called the National Evaluation Council (CONEVAL). This council maintains close working relations with both Congress and the finance ministry. CONEVAL is in a position to develop a reputation as a producer of credible and useful information and thus ensure that the M&E outputs it supplies are actually used for budgetary decisions in both branches of government.

*Source: Compiled by author.*
their own evaluation plans and had to ensure that the entire ministerial portfolio would be evaluated every three to five years. The finance ministry played a key role in quality assurance, but otherwise its main role in the evaluation system was to insist on the use of evaluation results during budget negotiations. Internal surveys carried out by the Australian government in the 1990s found evaluation results to be very influential, particularly in deliberations over new policy initiatives (Mackay 2004).

In weaker administrative environments, it might not be possible to delegate the operation of M&E instruments because of lack of capacity, or because it might compromise their credibility. However, even if the public administration is weak or fragmented, the ultimate measure of success for an M&E system is not the extent to which the ministry of finance operates every aspect of it: M&E is only effective when budget analysts, legislators, and political leaders use the M&E data produced to formulate and negotiate the budget. There are several key factors that can severely inhibit the use of M&E for budgetary purposes. A government needs reliable information systems, solid financial management systems, and at least the basic human capacity to generate and process evaluations. It does not, however, need a very systematic and planned M&E system to start with, especially not in countries with lower capacity. Uganda is an example of a country where the creation of a new budget directorate, out of the 1992 merger of the planning and finance ministries, created a big enough unit of professionals to conduct high-level analytical work even though much of the public sector was lagging far behind. Such islands of excellence can make use of performance information from many sources, including from donors, and try to spread their results without relying on a demanding framework that systematically links performance data to the budget process (Kuteesa and others 2009).

**How Best to Implement M&E-Budget Links**

Institutional change rarely happens because the benefits of change are self-evident. In budgeting, large amounts of money and quite often political futures are at stake, and as long as the budget process works reasonably well, budgetary actors are understandably reluctant to change the institutional setup. Sources of potential resistance against M&E include the finance ministry itself, particularly the budget office. In many finance ministries, budget analysts are very skilled at challenging ministerial budget proposals and successfully trimming down excessive spending. Especially when the introduction of performance information goes hand in hand with the dismantling of traditional line item controls to “let managers manage”; budget analysts might see more danger than gain in reform.

In most OECD countries where finance ministries did take the lead and successfully implemented M&E-oriented reforms, a few factors were usually present:

- Top officials in the finance ministry were convinced that the present budgetary toolkit was not sufficient for current or looming fiscal challenges. Periods of fiscal austerity, volatile revenues, or crises of confidence in the government’s ability to spend wisely often combined to create such a sense of urgency.
- There was an individual or a group of institutional champions to “sell” both the technical and the political benefits of reform to different groups of stakeholders.
- Finally, the first two factors would eventually combine to make it feasible for ministers and other political leaders to expend the political capital and attention necessary to overcome resistance and adopt the necessary decisions.

Prospective reformers don’t have to wait for the next recession and a favorable political climate to attempt setting up an M&E system. In many countries, different budgetary actors successfully experimented with different tools and instruments before the reform climate reached a tipping point and piecemeal change turned into comprehensive reform. In Mexico, the Ministry of Social Development had been working for years on performance indicators, logical frameworks, and evaluations before the finance ministry started to champion reforms on a larger scale. In Chile, reforms accelerated markedly after the country was hit by the effects of the Asian financial crisis. The success of today’s comprehensive system benefitted greatly from the experience of a first generation of reforms that some champions within the budget office had implemented. Australia, Sweden, the Netherlands, and the United Kingdom are cases where changes in government prompted dramatic changes over very short peri-
ods of time, which were then often followed by many years of fine-tuning.

There does not seem to be a golden rule as to how exactly an M&E function should be anchored within the ministry of finance. There are several cases, including Mexico and Chile, where ministers chose to set up a dedicated M&E unit within the budget office separate from budget analysts. Such a setup ensures that M&E has internal champions charged with promoting the new instruments, yet at the same time it also creates an institutional barrier between M&E and the officials who are in charge of drafting the budget. On the other hand, at the UK Treasury, as well as at the U.S. Office of Management and Budget (OMB), the budget analysts are directly in charge of managing and processing M&E information. OMB’s Program Assessment Rating Tool (PART) was managed by its individual budget examiners and has been linked closely to the preparation of the executive draft budget. In some cases, officials have noted that such a setup slows down the first implementation of M&E, because budget analysts are so busy running the budget that they find little time to take charge of the new instruments.

Inevitably, the right setup will depend on the specific administrative context. The available evidence, in particular from the United Kingdom, encouragingly suggests that the longer performance budgeting is in use, the more support there is among budget officials. The support seems to grow not only because M&E turn out to be useful as budgetary instruments, but also because budget analysts come to see the work with M&E as more policy oriented and thus more relevant and attractive than the traditional handling of line items.

Key Lessons

There are four key lessons:

• **There is no model M&E system that works in every country.** Successful introduction of M&E into the budget process has to account for the existing budgetary framework, and very different models have been successful in different countries.

• **M&E is a demand-driven process.** Even the most sophisticated M&E instruments are ultimately ineffective unless they are actively used by public officials for budgetary decisions.

• **Introducing M&E is a joint effort.** When setting up an M&E system, officials in the finance ministry cannot work in isolation, they need support from ministries and agencies and interest from legislators.

• **To bring M&E into the budget is a politically sensitive reform.** To use M&E properly for budgeting, established administrative practices need to change, which will require a sustained reform effort and long-term political backing.

About the Author

Philipp Krause specializes in public administration and budgeting and is completing his PhD at the London School of Economics and Political Science. He is currently a visiting scholar at New York University’s Wagner School. Philipp has previously worked on public sector issues for the German Technical Cooperation and the World Bank, and with governments in Latin America, Africa, Europe, and the Middle East. He has written articles and reports on public sector reform, budgeting, and fiscal governance.

Endnotes

1. The author thanks Helena Hwang (Consultant, Poverty Reduction and Equity Group), Mariano Lafuente (Research Analyst, Latin America and Caribbean Public Sector Governance), Nick Manning (Adviser, Public Sector Governance), Gladys Lopez-Acevedo (Senior Economist, Poverty Reduction and Equity Group), Keith Mackay (Consultant, Poverty Reduction and Equity Group), and Jaime Saavedra (Acting Sector Director, Poverty Reduction and Equity Group) for their very helpful comments. The views expressed in this note are those of the author. To access other notes in this series, visit www.worldbank.org/poverty/nutsandbolts.

2. For current information on the World Bank’s ongoing work on performance budgeting, please refer to the Public Sector Governance Web site: http://go.worldbank.org/SGO4LFRSS0.

3. Among other examples, Hood (2006) reports a UK experience: hospital patients waiting in lines of ambulances outside emergency rooms to meet a target for patients to be seen by a doctor within a certain time of admission.
Further Reading


This note series is intended to summarize good practices and key policy findings on PREM-related topics. The views expressed in the notes are those of the authors and do not necessarily reflect those of the World Bank. PREM notes are widely distributed to Bank staff and are also available on the PREM Website (http://www.worldbank.org/prem). If you are interested in writing a PREM note, email your idea to Madiguene Seck at mseck@worldbank.org. For additional copies of this PREM note please contact the PREM Advisory Service at x87736.

This series is for both external and internal dissemination