CLEAR Training Materials

Performance-based Budgeting

Manual

clear

Regional Centers for
Learning on Evaluation and Results

CLEAR (Centers for Learning on Evaluation and Results) is a global initiative aimed at strengthening developing countries' capacities in monitoring and evaluation (M&E) and performance management (PM) to support a focus on results and evidence-based decision-making. Working with regional institutions, CLEAR responds to increasing government and civil society demands for practical and applied M&E and PM capacity-building and to the current limited availability of relevant programs in many developing countries in several regions.

CLEAR is being supported by the African Development Bank, the Australian Agency for International Development, the Asian Development Bank, the Belgian Development Cooperation, the Inter-American Development Bank, the Rockefeller Foundation, the Swedish International Development Agency, the Swiss Agency for Development Cooperation, the UK Department for International Development, and the World Bank Group. The secretariat is housed in the Independent Evaluation Group (IEG) of the World Bank Group.

More information on CLEAR is available at www.theclearinitiative.org.

Performance Budgeting

Marc Robinson

Contents

Table of Acronyms

AA	Accrual accounting
ABC	Activity based costing
AOB	Accrual output budgeting
BC	Budget classification
COA	Chart of accounts
COFOG	Classification of the Functions of Government
DRG	Diagnostic related group
FF	Formula funding
FMIS	Financial management information system
GBE	Government business enterprise
IFMIS	Integrated financial management information system
IT	Information technology
LOLF	Loi organique relative aux lois de finances
MFR	Managing for results
MoF	Ministry of finance
MT	Medium term
MTEF	Medium-term expenditure framework
MTFF	Medium-term fiscal framework
OECD	Organization for Economic Cooperation and Development
PB	Performance budgeting
PFM	Public financial management
PM	Performance measurement
PSA	Public Service Agreement
SAI	Supreme audit institution

ZBB	Zero-base budgeting

Introduction

This manual is designed to accompany the courses on performance budgeting delivered at the regional Centers for Learning on Evaluation and Results (CLEAR). It presents summaries of the essential content for each of the core topics covered in those courses. This manual can therefore be used on a "stand alone" basis. However, it may be used to best advantage in conjunction with the support of the key readings listed at the end of each section.

1. Budgeting and the Budget Process

Government budgeting is the allocation and use of resources, and associated decisions about how the resources used will be acquired, by that part of the public sector which is financed primarily by compulsory charges such as taxes [1]. Equivalently, government budgeting – "budgeting" for short in what follows — can be said to be about the allocation of resources for the provision of services and transfers on a non-market basis, where "non-market" means either free or greatly below cost.

The government budget is therefore a financial plan which covers those public sector agencies – such as ministries and the social security system – which are primarily tax-financed. It does not typically cover public corporations, which normally manage their finances on a more autonomous basis. [2]

Budgeting is a process involving four stages:

Aggregate fiscal policy formulation: this refers to the determination of the government's overarching objectives for the budget deficit, debt and other relevant fiscal aggregates [3], which should then be translated into decisions about the desired levels of aggregate revenue and expenditure.

Budget preparation and enactment: in this stage of the budget process the government decides how much funding it will provide to which agencies and for which purposes. This is given formal expression in the budget law and budget regulations, which are enacted by the legislature and the highest executive organs of government (i.e. the president or cabinet of ministers).

Budget execution: this refers to the carrying out of the expenditure plan developed in the budget – including the entering of contracts and expenditure of funds.

[1] We use the term "tax" as shorthand to include other compulsory revenue sources such as mandatory levies or fees, and fines. Budget financing via borrowing may be considered to be financing by deferred taxes.
[2] Public corporations are financed primarily by the prices which they charge the consumers of their services – prices which are voluntary in that consumers are not legally obliged to purchase the services concerned. The budget normally only relates to public corporations insofar as they either receive taxpayer subsidies or their profits are a supplementary source of budget funding.
[3] Depending on the policy framework of the country concerned, the other fiscal aggregates in terms of which fiscal policy is framed might include, e.g., the level of "net worth" (government assets minus liabilities) or the percentage of revenue/GDP.

Accounting, Auditing and Reporting: this covers the preparation of accounting records of government spending and revenue, their auditing by both internal and external auditors, and the provision of reports on budget execution to government agencies, ministers, parliament and the public.

Budgeting has three key objectives

- *Sound macro-fiscal outcomes*: this is primarily the product of good aggregate fiscal policy. Aggregate fiscal policy needs to ensure "fiscal sustainability" – which means essentially that deficits and debt do not get out of control – because this is important for economic confidence and stability. If budgeting does not keep debt to sustainable limits, financial markets can potentially lose confidence in the security of lending to government. The result may be a sudden increase in the interest rates which government is forced to pay and, in the extreme, an unwillingness to lend to government. In addition to fiscal sustainability, aggregate fiscal policy can play an important "stabilization policy" role – in particular, through providing support to the economy through additional spending or lower taxes during a recession.

- *Appropriate prioritization of expenditure*: this refers to the allocation of funds to the sectors and programs which are most effective in meeting social needs. It means, for example, that if a country is facing a major new health challenge, additional funds will be allocated in the budget for programs designed to tackle that challenge. The corollary of this is that funds are moved away from sectors and programs where spending is low priority.

- *Service effectiveness and efficiency*: this means, firstly, that government services are delivered *efficiently* and, secondly, that they are designed and managed so as to maximize their effectiveness.

Performance budgeting, as we will see, focuses directly on improving the ability of budgeting in to deliver on the second and third of these objectives. [4] "Effectiveness" and "efficiency" are therefore key words here. Effectiveness refers to the extent to which a service delivers the benefits which it is supposed to deliver to society. For example, an HIV/AIDS prevention campaign is effective if it succeeds in reducing the rate of new HIV/AIDS infections. As will be explained later, effectiveness is about the extent to which programs achieve their intended *outcomes*. Efficiency, on the other hand, refers to delivery of services at the lowest possible

[4] However, in doing so, it can also indirectly improve macro-fiscal outcomes.

cost, without sacrificing quality. Expressed differently, efficiency is about minimizing waste in the production of government services.

Budget preparation should start with, and be carried out within the context of, good aggregate fiscal policy. In other words, decisions about how much funding to provide agencies should be constrained by clear policy about the level of aggregate expenditure which is, given government revenues, consistent with government's deficit and debt objectives. Key aspects of budget preparation and enactment include:

- The determination of expenditure priorities,
- Estimates of spending requirements – particularly associated with ongoing commitments such as the government wage bill, capital projects underway, and expenditure obligations on "mandatory" items such as pensions and other social security payments,
- The formulation of current and capital budgets, which require different budget preparation methods but which should at the same time be linked. [5]
- The formulation of budget funding decisions into legal authorizations ("control totals") which can be "appropriated" by the parliament or promulgated by the executive government.

The way in which budget preparation is organized differ considerably between different countries. Some key aspects of budget preparation which have a particularly close bearing on performance budgeting – such as expenditure prioritization and medium-term budgeting methods – are discussed in this manual.

Budget execution, on the other hand, can be seen as a clearly defined sequence of stages, as follows:

- *Authorization stage*: based on the parliamentary budget appropriations, the finance ministry communicates to ministries their expenditure entitlements. The nature of these authorizations differs internationally. In some countries, authorizations are made on a month-by-month basis – that is, each month ministries are told how much money they will have to spend in the coming month.

[5] Building major new capital assets such as schools, hospitals and roads has important implications for future current expenditure, whether on staff and supplies to run the facilities concerned, or on the maintenance of those assets.

- *Commitment stage*: a commitment is a contract or some other obligation to make a future payment.
- *Verification stage*: ensuring that the service or goods previously contracted for have been delivered, and meet the specifications of the contract.
- *Payment authorization stage*: in the light of verification, approval for the payment to be made is given by an official independent of the persons who initiated the commitment.
- *Payment stage*: the payment is actually made.
- *Accounting stage*: each payment is entered into the accounting system and its relevant characteristics recorded. In a well-developed system, the accounting system will also record the commitments stage.
- *Auditing stage*: the accounts are audited – firstly by "internal auditors" (within the spending ministry concerned) and subsequently by "external auditors" (the auditor-general or audit court).

Performance budgeting aims to change budget preparation fundamentally, linking it much more systematically to the effectiveness and efficiency of expenditure. In doing so, however, it also brings about equivalently major changes in budget execution.

Key Readings

Potter, B. and J. Diamond (1999), *Guidelines for Public Expenditure Management* (Washington: IMF), available at http://www.imf.org/external/pubs/ft/expend/index.htm.

Schiavo-Campo, S. & D. Tommasi (1999), *Managing government expenditure* (Manila : Asian Development Bank), obtainable at http://www.adb.org/documents/manuals/govt_expenditure/.

2. Overview of Performance Budgeting

Performance budgeting aims to improve the effectiveness and efficiency of public expenditure by linking the funding of public sector organizations to the results they deliver. It uses systematic performance information (indicators, evaluations, program costings etc) to make this link. The impact of performance budgeting may be felt in improved prioritization of expenditure, and in improved service effectiveness and/or efficiency.

PB usually also emphasizes giving government agencies and their managers greater flexibility in the use of resources than they would typically have under traditional tightly-controlled public management systems. A key element of this is greater flexibility in the choice of the mix of inputs which are to be used to deliver services (e.g. how much labor input vs. externally sourced inputs, the mix of types of externally-source supplies and services used). An important implication of this is the need for more flexibility human resources management, a topic discussed in a later section.

The increased international interest in performance budgeting has been prompted in part by a recognition that it is all too easy in Government to lose sight of the fundamental objective of delivering positive outcomes to the community. Public sector organizations which are financed by taxes and other compulsory charges lack the market disciplines which compel commercial enterprises, particularly those operating in highly competitive markets, to be customer-oriented. Political accountability through the electoral process is, of course, extremely important, but is not necessarily sufficient to ensure that public sector organizations are highly focused upon the results they deliver.

Expenditure Prioritization

Expenditure prioritization is often quite weak in government. A particular concern is the tendency for much ongoing program expenditure to escape serious scrutiny, and for budgetary decisions to be mainly focused on what new spending initiatives to adopt. The term "incrementalism" is widely used to describe this tendency. [6]

If prioritization processes are not well developed, governments typically rely, if and when they need to make spending cuts, on "across-the-board" cuts – that is, cuts of the same percentage

[6] For more on incrementalism, see Robinson and Brumby (2005).

to all ministries. Across-the-board cuts are by definition indiscriminate and inferior to a selective approach which would target cuts at lower-priority or less effective programs.

Improved expenditure prioritization can help to achieve more sustainable public finances. If, for example, spending cuts are needed to achieve "fiscal consolidation" (i.e. to get deficits and debt under control), the capacity to target cuts selectively at the lowest priority programs will help to make those cuts more sustainable (i.e. the reduce the pressures to reverse them), thus increasing the probability that the improvement to the overall state of public finances will be enduring.

Managing-for-Results

Performance budgeting should be viewed in the broader context of a set of related "managing-for-results" (MFR) reforms. MFR can be defined as the use of formal performance information to improve public sector efficiency and effectiveness. Its fundamental starting point is maximum clarity about the outcomes which government is attempting to achieve, and about the relationship of outputs, activities and resources used to those desired outcomes. Good strategic planning and business planning are an essential element of MFR. MFR also tends to emphasize the *ex ante* stipulation of performance expectations for agencies, work units and individuals through the use of performance targets and standards.

A standard element of the "strategic human resources management" component of MFR is the introduction of stronger performance-based extrinsic incentives (rewards and sanctions) for public officials. Typically, this is accompanied by greater flexibility of employment, including greater capacity to sanction or dismiss poor performers, and greater ease in transferring or terminate employees in programs which the government is eliminating or cutting back.

Examples of other elements of MFR reforms which have been introduced in many countries include:

- Customer orientation measures: an example of such a measure is "client rights charters", which define client service rights (e.g. timeliness of service, what type of service etc).

- Market-type reforms: this refers to a range of measures whereby public service providers are made to operate in a more business-like manner, under the pressure of certain types of market forces or mechanisms which simulate markets. Competitive tendering – where a

service provided by a government agency is thrown open to private competition via tender – is one example. Purchaser-provider systems are a performance budgeting model which falls into the category of market-type reform.

Models of Performance Budgeting

There is not just one model of performance budgeting, but a range of different models. All link funding and results, but in different ways.

Some performance budgeting systems – program budgeting and zero-base budgeting, for example – are government-wide systems. On the other hand, some are only intended to apply to particular sectors of expenditure or categories of organization (e.g. a formula funding system applying to schools).

Performance budgeting systems differ to some extent in their objectives. Some place aim principally to improve expenditure prioritization, while others are mainly focused on in improving service effectiveness and/or efficiency. Reflecting the different forms of linkage between results and funding which they seek to build, performance budgeting systems also differ in respect to the type of performance information upon which they primarily rely. This applies both to the type of performance measures which they use, and also to whether they make use of evaluation as well as indicators. As discussed earlier, all forms of performance budgeting have in common the idea of greater managerial freedom, particularly in respect to the choice of input mix.

The most basic form of performance budgeting is that which uses performance information systematically in the preparation of the government-wide budget. A common tool used for this purpose is *program budgeting*, in which expenditure is classified in the budget by objectives (outcomes and outputs), rather than solely by economic categories (such as salaries, supplies and communication costs) and organizational categories (e.g. ministry and department with the ministry). The primary objective of program budgeting is improved expenditure prioritization. However, by increasing the attention paid during budget preparation to spending ministry

performance, this type of performance budgeting also aims to increase the pressure on ministries to improve efficiency and effectiveness. [7]

It should be emphasized that, in the terminology adopted here, performance budgeting is a generic concept of which program budgeting is one form of performance budgeting. This needs to be emphasized because there are some who attempt to distinguish program budgeting from performance budgeting.

Zero-base budgeting (ZBB) is essentially a variation of program budgeting which, in its original form, called for the comprehensive review and prioritization of all expenditure on a continuing basis. To achieve this, all programs were to be decomposed for the purposes of budget preparation into "decision packages" (also known as "service increments") which would provide choice about the extent to which each program might be cut back or, alternatively, given increased funding. Ideally, these decision packages would cover all major options, even to a 100 percent cut in the program concerned (hence the "zero base"). Priority rankings would be attached to these decision packages – again, on the basis of performance information – and these rankings would then be used to ensure that the available level of revenue funded those decision packages which were of highest priority. In this way, ZBB aimed to go even further than program budgeting in improving expenditure prioritization.

The fundamental problem with ZBB was the practical impossibility of comprehensively reviewing *all* expenditure each year. In other words, the "zero base" examination of each program was not really practicable. Selective expenditure appraisal is all that, in practice, can be achieved. In recognition of this, ZBB evolved into what is sometimes call *alternative budgeting.* Rather than seeking to examine all options for each program, extending even to a one-hundred percent cut, alternative budgeting tried to make the task manageable by confining the process to the examination each year of a much narrower range of options for cuts or increases to each program – for example, the options of 15, 10 and 5 percent cuts, and perhaps also of 5 or 10 percent increases.

[7] Program budgeting has the potential to contribute significantly to aggregate expenditure discipline, and thereby to fiscal discipline. By giving government an improved capacity to identify low priority or ineffective programs which can be cut, it makes it easier to create the fiscal space necessary to respond to new policy priorities without undue growth in aggregate spending. Moreover, if and when fiscal consolidation is necessary, better prioritization makes it possible to target spending cuts so as to minimize their social cost and increase their sustainability.

Under both program budgeting and zero-base budgeting, the link between results and funding is a rather loose, rather than mechanical, one. The results achieved by programs (or decision packages with programs) is considered closely in the budget preparation process, but there it is not the case that poor results automatically means cuts – or, conversely, that good results automatically means increasing funding.

Since the 1980s, a range of newer forms of performance budgeting have been developed which aim to build a tighter linkage between funding and results. These newer forms of performance budgeting all aim to increase the pressure on public agencies to improve the effectiveness and/or efficiency of the services which they delivery. To achieve these, they rely on three mechanisms, used either in isolation or in combination.

The first of these mechanisms is *linking budgets to performance targets.*

Setting performance targets – whether for ministries, work units or individuals – is, as noted above, a general MFR theme. It is when budget funding is linked in some manner to performance targets that we are talking about a form of performance budgeting. The best example of a target-based form of performance budgeting is the *UK Public Service Agreement* (PSA) system as it operated between 1998 and 2004 [8]. Under this system, targets were set on a triennial basis as part of a multi-year budgeting framework. In other words, high-level targets for outcomes and outputs (see the next section for explanations of these key concepts) were set as an integral part of the process which, every three years, determined the core funding levels which each ministry would receive for the coming three year period. Targets were set in the light of funding, and funding in the light of performance against targets. The major question that this raises – to be discussed later in section 13 – is the *concrete* nature of this link between funding and targets.

Target-setting raises a range of issues, which will also be discussed in detail later section 13. One of these is the problem of selecting the right numerical value for the target. In principle, of course, the target should be neither too tough nor too easy. However, it is often not easy for the ministry of finance (MoF) or other central decisions makers, without a detailed knowledge of the area of service delivery concerned, to set the target at an appropriate level. Another issues raises by target-setting is the issue of behavioral distortions, an example of which is the danger that, in striving to fulfill a target set in terms of one dimension of performance, the agency may

[8] See P. Smith, "Performance Budgeting in England" in Robinson, *Performance Budgeting.*

put less effort into another dimension of performance which is not factored into the performance target. [9]

The second mechanism used to creating a tighter link between funding and results is *formula funding*, which is the subject of section 12. Formula funding means that the level of funding to a government agency is determined (in part or whole) as a mathematical function of some explicit variables. A simple example is a formula to estimate school funding requirements over the medium term as a function of the number of students and the cost per student (based on demographic projections). This illustrates the idea of a *cost-based* formula, the simplest version of which bases funding on the quantity of output times the unit (average) cost per output – as, for example, in a school funding system where funding for each school is determined primarily by its number of students multiplied by an amount intended to cover the cost of one year's education of one student at the relevant stage of schooling (e.g. $4000 per year nine student per year).. In this case, the formula is being used only as a budget estimation tool – that is, to estimate how much funding should be provided to a particular service area or to specific service provider units (e.g. specific schools).

Not all formula funding arrangements can be said to be forms of performance budgeting. It is only when the formula links funding to results (outcomes and/or outputs), and does so with the intention of boosting efficiency and/or effectiveness, that formula funding becomes a tool for performance budgeting. This would be the case, for example, if the formula was being used in effect to set a performance target for the agency. As a concrete example of this, suppose a ministry was funded on a formula basis to, say, vaccinate one million persons at $20 per vaccination in the coming year, and was then held to account for if it spend all the money but only vaccinated 600,000 people.

The third mechanism used by the newer performance budgeting models is *performance funding incentives* to agencies. An example of this is performance bonus funding, which operates by giving agencies additional payments for measured good performance (or perhaps also funding reductions as a function of measured poor performance). An example is university funding systems which operate in many countries in which public universities are given bonus funding, over and above their core funding, based on performance against measures such as the percentage of graduates who are in professional employment six months after graduation. As

[9] For example, if the target is set in term of the unit cost of service delivery, the agency might sacrifice quality in order to cuts costs.

this example shows, most performance funding incentives are also forms of formula funding, because the funding incentive provided is determined by a mathematical formula linked to a performance indicator. [10]

Performance funding incentives – like formula funding – are almost invariably used as instrument for *sectoral* performance budgeting systems (i.e. covering some sector of government, such as schools, or hospitals.

The above outlines three key mechanisms used by newer performance budgeting systems to create tighter links between funding and results. As indicated above, these mechanisms may be used in isolation or in combination. An important example of the combined application of whoe of these mechanisms is the *purchaser-provider* (PP) model. PP is a combination of the principles of incentive payments and formula funding (the second and third of the mechanisms identified). In a purchaser-provider system, a public sector agency is funded as if it were an arms-length provider of goods or services. Government funds the agency by paying a "price" per unit for the agency's outputs (and sometimes, but more rarely, outcomes). If the agency's cost of production exceeds the price it is paid, the agency makes a loss. Conversely, it retains any profit made by producing at a cost below the price. In a PP system, funding is only provided for results delivered. Finally, the price paid is ideally based on, partly or in whole, on some measure of what it should cost an efficient producer to deliver the output concerned. The aim of this system is to motivate strong performance via strong financial performance incentives.

The most successful example of a purchaser-provider system is the "diagnostic-related group" system of hospital funding. Broadly speaking, this is a system in which hospital are funded primarily for the treatments they deliver to patients. Services are categorized into a number of different output types ("DRGs"), and a specific price is attached to each DRG category. Such a system provides a powerful incentive for efficiency. At the same time, it raises concerns about the danger of cost-cutting which undermines the quality of treatment or has other perverse behavioral effects. The experience of the DRG system in this respect is, however, encouraging.

Accrual output budgeting (AOB) represented an (unsuccessful) attempt, in New Zealand and Australia, to apply the purchaser-provider principle to the government as whole, as opposed to a particular sector such as hospitals. The idea of AOB was that ministries would be funded not

[10] Note that in this case, the formula is not linked to the costs of the performance delivered (for example, no-one would pretend to know the additional cost involved in boosting the graduate employment rate by, say, 5 percent).

by being given budgets, but would rather be paid for the outputs they deliver, as if they were external contractors to government. It was intended that government would be a tough price-setter, seeking to create the type of pricing pressure which occurs in a competitive market. Because in this model ministries were supposed to act like businesses, it was necessary that they use business (i.e. accrual) accounting – hence the "accrual" in "accrual output budgeting" (the link with accrual accounting is discussed in a section 16). It also became necessary to distinguish government's "purchaser" role vis-à-vis its agencies from its "ownership" role. The AOB experience is discussed in detail in the section 12.

Key Readings

Robinson, M (2007), "Performance Budgeting Models and Mechanisms" in M. Robinson (ed.) *Performance Budgeting: Linking Funding and Results*.

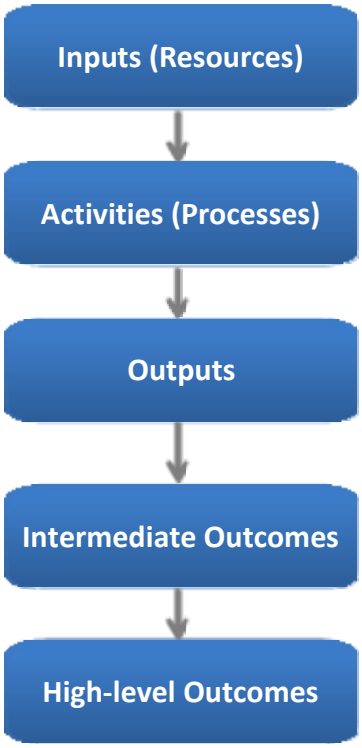
3. Performance Information Fundamentals

"Outcomes" and "outputs" play a central role in all models of performance budgeting, and it is essential for any discussion of performance budgeting that these and related concepts are clearly understood.

Performance Concepts: the Results Chain

In the *results chain* framework, outputs are produced using inputs (resources) via activities and processes, and outputs generate outcomes for the community.

The Results Chain



Outputs are goods or services – the "products" – which a ministry or other government organization delivers to external parties. This usually means services delivered to or for the direct benefit of the community, although some outputs are services which ministries deliver to other ministries. Examples of outputs include: medical treatments; advice received by farmers from agricultural extension officers; students taught; and police criminal investigations. Most

government outputs are services, and for this reason the terms output and service will be used interchangeably in this manual.

Outcomes are the intended impacts of outputs – more precisely, the changes brought about by public programs upon individuals, social structures, or the physical environment. Health inspections of restaurants are an output, the intended outcome of which is that fewer diners fall sick. Criminal investigations are a police output, and reduced crime the outcome.

Many government services aim to achieve more than one outcome. For example, school education aims to increase the level of education of the population. But it also aims, amongst other things, to improve economic performance. Both a higher level of education and a stronger economy are outcomes. Because it is by means of the first of these that the second is achieved, a more educated population is said to be an *intermediate outcome*, and a stronger economy a *higher-level outcome*. The relation between proximate and high-level outcomes is one of logical causality (i.e. the proximate outcomes induce the high-level outcomes).

The outcomes element is the distinctively public sector component of the results chain. By contrast, the way in which outputs are produced is conceptualized in the results chain in exactly the same way as in the private sector: outputs are produced by the use of *inputs* in production *activities and processes*. For example, the treatment which seriously injured person receives in hospital involves the use of a set of inputs (skilled staff, operating equipment and facilities, medical supplies, electricity etc) and a set of activities including anesthesia, surgery and nursing, as well as supporting activities such as supplies and facility management.

Inputs, as this example indicates, refer to all inputs, assets and capabilities which are or may be drawn on in the production process to deliver the outputs and outcomes desired. Although "inputs" is the conventional results chain term, and therefore will be used here, the term "resources" actually captures better the scope of what is referred to. Thus inputs which contribute to the *capability* to deliver results include not only equipment and buildings by, for example, organizational culture and staff morale. [11]

[11] We are, in other words, talking about much more than the "factors of production" of economics 101 (labor, land and capital), and also to draw attention that the inputs an organization can drawn upon are not only purchased from outside, but created within the organization.

Physical assets are inputs which are sometimes mistakenly thought of as outputs – as in, for example, the number of roads or bridges provided by government to the community. But it is not the roads and bridges which are outputs, but rather the service which citizens get from these roads and bridges. This means, for example, that passenger miles traveled is an output measure while kilometers of road is an input measure. [12]

The term *activities* may seem self-explanatory, but confusion between activities and outputs is very common. Some examples can help avoid this confusion:

- In a hospital, anesthesia and cleaning are activities rather than outputs because they are components of the overall service provided to the patient, rather than the complete service. The patient can't recover through anesthesia or cleaning in isolation, and it is only via the combination of all the necessary activities that the complete service (the output) is delivered. More generally, *an output must be capable of delivering the intended outcome*.

- Bus driving is an activity, whereas passenger trips are the outputs. Similarly, teaching is an activity rather than an output, and the output is students taught. In these cases, an activity produces multiple outputs. [13]

As mentioned above, outputs are defined as services provided by ministries or other government agencies to *external* parties. So if one ministry provides services to another ministry – for example, training, or office accommodation – these are ministry outputs because the other ministry is a client *external to the ministry providing the service.* However, a service which one part of a ministry provides to another part of the same ministry – for example, when the education ministry's human resources group recruits teachers for the primary school group – is not an output but a *support service.*[14]

[12] This should be qualified by noting that, if a separate public works ministry is responsible for constructing the roads and bridges and then delivering them to the transport ministry which manages them, then the roads and bridges would indeed constitute outputs *of the public works ministry* – because it is delivering them to an external client. But if the transport ministry manages the construction itself, the roads and bridges are not outputs. In either case, the output provided by the transport ministry is the use of the road by travelers.

[13] Or, potentially, as illustrated by the case of bus driving, no outputs at all. This would be the case if the bus ran but no passengers got on board.

[14] Some would argue that what constitutes an output, as opposed to a support service, depends on your organizational perspective, and that from the point of view of the human resources group, recruitment is an output which they deliver. This is a valid observation. However, in the performance budgeting context, the perspective normally adopted is that of the ministry or other distinct government organization. The point is to focus entire ministries on the services they deliver to external clients, rather than on internal activities.

The results chain is a very well-established conceptual framework for the public sector which goes under a range of names, including the "program logic model" and "logical model". Although its concepts are standardized, and have clear meanings attached to them, there is one significant divergence of usage worth mentioning: the practice adopted by some of dividing outcomes, as defined above, into "outcomes" and "impacts". The distinction is either one of timelines (impacts being more long-term), or is the distinction between intermediate and high-level outcomes. This terminology is not used here, because in practice there is often no clear distinction at the margins between outcomes and impacts.

One other minor point of terminology: "intermediate" and "high-level" outcomes go by a wide range of other names (immediate outcomes, proximate outcomes, end outcomes, ultimate outcomes). The idea is, however, the same.

Outcomes vs. External Factors

Outcomes are the *changes brought about by government intervention*. If the level of malaria falls for reasons which have nothing to do with government actions – because, for example, there is a drought which reduces mosquito numbers substantially – this fall is not a government outcome. Neither is it an outcome if the rate of economic growth increases substantially because, and only because, the world economy is very buoyant. The fall in malaria, or the boost in the growth rate, are outcomes only to the extent that they are the result of government actions.

Outcomes therefore need to be distinguished from the consequences of *external factors*. External factors are factors beyond the control of government which influence the characteristics of individuals, social structures or the physical environment which the government is trying to change. The level of rain is therefore an external factor impact impacting on the malaria rate. The state of the world economy is an external factor impacting on the rate of domestic economic growth. In assessing outcomes, the challenge is to distinguish the impact of external factors from that of the government intervention.

External factors are also sometimes known as *contextual factors* or *confounding factors.* External factor may be part of the external environment in which the program operates, or they may be characteristics of the client. An example of the former is the impact of the state of the economy on the success rate of a program for the rehabilitation of injured workers in getting its

clients back into the workforce. An example of the client characteristic type of external factors can be seen in education, where the educational levels attained by children are determined not only by the quality of the education they receive at school, but by external factors such as their intelligence, nutrition and level of parental support. Expressed differently, these types of external factors impact upon "outcome variables" such as the literacy and numeracy rates (see the next section for the concept of an outcome variable). This makes it difficult to compare the performance of two different schools by simply looking at levels of student educational attainment, because differences in the socio-economic mix of student populations can make such comparisons completely misleading.

Outputs vs. Outcomes and Activities

It is useful to clarify further the relationship between outputs and outcomes. To be considered an output, an output must be capable of achieving its intended outcome, but this does not necessarily mean that each unit of output will be successful in actually achieving the outcome. Consider the example of a medical treatment of an accident victim who dies notwithstanding receiving exactly the best treatment in a timely manner. It would not be reasonable to say under these circumstances that the patient did not receive a service (output). The same point applies, say, to education, where despite receiving excellent teaching, certain students may fail to learn much, while others in the case will do very well.

An associated point is that, to be an output, a service must be complete. Consider the example of a car, which is only a product (i.e. an output) only when all the parts (wheels, engine etc) are attached and it can therefore function as intended. Another good example is tetanus injections, where it is necessary to have a full course (of three injections) in order to acquire immunity. The tetanus treatment outputs should therefore be measured in terms of numbers of persons who receive the complete course. Otherwise, one will end up counting activities (e.g. incomplete course of tetanus treatments) which are incapable of yielding any benefit to the client/community.

Performance Indicators

Performance information refers to information on results achieved, and/or costs of achieving those results. There are two basic types of performance information: performance measures and evaluation. In this and the next couple of sections, we focus on measures. After that,

evaluation will be discussed. It should, however, be emphasized at the very outset that it is a mistake to think that performance budgeting is only about the use of performance *measures* in the budget process.

Performance indicators are quantitative measures which provide information on the effectiveness and efficiency of programs and organizations. There is no difference between a "performance indicator" and a "performance measure" – both terms are used interchangeably in this manual. We also need to be careful not to confuse objectives, indicators and targets. An *objective* is a statement of what one is trying to achieve – for example "reducing death from HIV/AIDS". By contrast, a performance *indicator* is quantified (e.g. "the percentage of the population which is HIV/AIDS positive", or "the number of persons dying annually from HIV/AIDS"). A *target* goes one step further and sets a precise aim to be achieved by a specific date (e.g. "reducing the percentage of HIV/AIDS-positive persons in the population by at least one-third by 2020").

Performance indicators should be selected according to the extent to which they are:

- Relevant
- Representative
- Cost-effective
- Comparable
- Minimize perverse effects

An indicator is *relevant* when the aspect of performance it seeks to measure is important to the objectives of its users. Relevance depends on who the user is. In selecting indicators to present to parliament and the public, for example, indicators of internal processes (such as time taken to fill job vacancies) would not be relevant.

An indicator is *representative* to the degree to which it succeeds in measuring the dimension of performance which it seeks to measure. Representative indicators are good proxies for what they are trying to measure. A measure of average time taken to answer client phone calls is, for example, a poor proxy for the quality of client service if what happens is that calls are answered quickly but the client is left on hold for an extended period. An indicator is more representative the less it varies for reasons which have nothing to do with performance, such as the impact of uncontrollable external factors and statistical uncertainty. To be representative, an indicator

should also be unambiguous. That is, it should be clear whether a change in the indicator means that performance has improved or deteriorated.

An indicator is *cost-effective* if the benefits of using a specific indicator exceed its cost of collection, processing and verification. Performance measurement systems are not cost-free. Like everything else, they should therefore pass a benefit/cost test.

A *comparable* indicator is one which is used by similar service providers, or by the same organization in the past. If, for example, there is an international standard measure for a service – for example, as with the measures provided in Europe by the common set of social policy indicators – then using that measure makes it possible to make performance comparisons with other countries. Similarly, performance comparisons become much easier if all relevant entities – e.g. all local governments, or all hospitals –a use the same indicators. Comparability over time is useful because continuing to use the same indicators that have been used in the past makes it possible to identify trends in before which might be impossible to detect if the performance measures have changed.

Perverse effects, which are discussed further below, are adverse consequences of performance measures. Examples are:

- In seeking to improve measured literacy and numeracy, schools might reduce their focus on other important dimensions of education, such as the teaching of foreign languages or social values,
- Suppose the timeliness of hospital accident and emergency ward treatment is measured by the percentage of patients who are treated within 4 hours. This can potentially produce the perverse effect that any patient who cannot be treated with 4 hours may be kept waiting considerably longer, with priority being given to patients who have arrived later but who can still be treated within 4 hours.

Perverse effects are in part the consequence of indicator design: for example, the omission of some key aspects of performance which, with improved design, could have been captured. In the case of hospital waiting times, the perverse effect referred to would be avoided by using a measure of *average* waiting time instead of the measure of those treated within a particular (arbitrary) time period.

These five criteria for good performance indicators are ideals, and few good indicators will score highly on all of them. Sometimes, a high rating on one criterion will come at the price of a

lower rating on another. Outcome measures such as the crime rate, for example, score very high on the relevance criteria but less well on representativeness because they are greatly influenced by "external factors" beyond the control of government.

In this section, we have focused on performance concepts and performance information in general. In the specific context of performance budgeting, however, what we are really interested in is the performance information – evaluations as well as indicators – which is useful for, and usable by, budget decision-makers. An important point to bear in mind here is that decision-makers can only absorb and use a limited and carefully-selected amount of information. It is therefore essential not only to be highly selective in choosing a handful of key performance indicators relevant to budget decision-makers, but also to put these indicators together with analysis which interprets and explains them, drawing overall conclusions about performance. This analysis should draw closely on the conclusions of formal evaluations and other review work.

Key Readings

Robinson, M (2007), "Informing Performance Budgeting" and "Results Information", in M. Robinson (ed.) *Performance Budgeting: Linking Funding and Results*.

HM Treasury et al (2001), *Choosing the Right Fabric – a Framework for Performance Information,* obtainable at http://archive.treasury.gov.uk/performance_info/fabric.html.

Royal Statistical Society (2003), *Performance Indicators: Good, Bad and Ugly*, obtainable at http://www.rss.org.uk/pdf/PerformanceMonitoringReport.pdf.

4. Performance Measures and the Budget

The indicators of most value for performance budgeting will tend to differ significantly from the indicators used for other purposes. What will generally be of greatest value to budget decision-makers in determining appropriate program funding levels will be indicators of the results achieved by programs – the outcomes that they achieve and the outputs which they deliver to achieve these outcomes. They will represent only a sub-set of the indicators used internally for managerial purposes by ministries, which will include not only results-oriented indicators, but also activity and input indicators focused on the internal processes, capacities and resources of the ministry.

There are three main types of output indicators: indicators of output quantity, indicators of output quality, and indicators of efficiency.

Output quantity indicators measure the volume of service provided. Examples of output quantity indicators are:

- Number of vaccinations carried out,
- Number of malaria prone districts sprayed,
- Number students taught at seventh grade,
- Number of planning applications determined.

Output quality indicators provide information on the extent to which the service is of a type which is likely to achieve its intended outcome. Two examples of more readily-measurable quality indicators are timeliness indicators and client satisfaction indicators. Examples of timeliness indicators are:

- Average waiting time of a hospital patient between arrival and treatment,
- Average time for a planning application to be determined,
- Average response time of the fire brigade to a fire.

Client satisfaction measures can be of various types, from simple measures of the level of satisfaction felt by the client, to more targeted measures such as:

- Client ratings of the courtesy of the service provider,
- The percentage of clients who consider that they obtained the service which they were seeking.

Efficiency indicators measure, ideally, the extent to which the service is delivered at low cost without sacrificing quality. Unit output cost one important type of efficiency indicator, although this is not a perfect measure as it does not necessarily hold quality constant. The unit cost is simply the total cost of delivering the output divided by output quantity. Examples of unit cost measures are:

- Cost per vaccination (including delivery),
- Cost per planning application determined,
- Cost per visa application processed.

Labor productivity indicators are another important type of efficiency indicator. Examples of labor productivity indicators include:

- Pension benefit applications processed per staff member,
- Average staff time taken to administer a practical driving license test.

In respect to outcome measurement, the biggest challenge arises from the distinction, discussed in the last section, between outcomes and external factors. As noted, outcomes are the change in characteristic of individuals, social structures or the physical environment brought about by government intervention, and the challenge is to separate out the impact of government intervention from that of other factors beyond government control.

In practice, many outcome indicators are in fact measures of what can be called the *outcome variable* —that is, the characteristics of individuals, society or the physical environment which public programs seek to change. Measures of the outcome variable do not distinguish the impact of the government's intervention from that of external factors. A good example of this type of outcome indicator is a crime rate statistic (e.g. the number of robberies per thousand of population). We know that the crime rate is not only influenced by government interventions such as policing, but is also greatly affected by social and economic developments beyond government control (e.g. the level of poverty). Crime rate statistics make no attempt to eliminate the impact of the latter type of external factor.

The perfect outcome measure would be one which excludes the impact of all external factors and which measures only the change in the outcome variable brought about by government intervention. Developing such measures is, however, exceedingly difficult. In a few areas, modern performance measurement techniques have gone quite a long way in developing outcome measures which remove, to at least some extent, the effects of external factors. A good example of this is the so-called *value-added* education performance measures, which attempt to adjust for student characteristics. For example, value-added school league tables present measures of the comparative performance of schools which adjust for these differences in student population. However, although this and other techniques have been developed to obtain better outcome indicators which are less effected by external factors, the overall scope for eliminating external factors from outcome measures is quite limited. As a result, the great majority of outcome measures are affected, to varying degrees, by external factors.

Ensuring that Indicators are Useful

To be useful to central decision-makers, who invariably have great demands on their limited time, program performance information needs to be readily digestible. A couple of key program indicators may, for example, be more useful to top decision-makers than a comprehensive compendium of dozens of indicators. The detail is of more interest to program managers than to the center.

A notable innovation in this context has been the development of summary measures which incorporate a wide range of performance information into one or more overall performance ratings for the program. A good example of this is the measures of program performance which was developed under the US Program Assessment Rating Tool (PART) which operated under the Bush presidency. Under the PART system, the US finance ministry (the Office of Management and Budget) set about rating all federal government programs over a period of 5 years. Each program was rated on a scale of 4 ratings, ranging from effective to not effective (there was also a "results not demonstrated" rating, used where there is insufficient information to form a judgment). These summary ratings were intended to be much more informative and readily understood than the large body of more detailed measures and evaluations which underpinned them. The program ratings, and the reasoning behind them, were all made public (on the website ExpectMore.gov). Moreover, the PART system was designed from the outset as a tool for performance budgeting, because the program ratings were used to inform the preparation

of the president's budget proposal to Congress (and were also intended to influence Congress' budget decisions).

Summary performance measures – of which PART is by no means the only recent international example – have the enormous advantage of ready comprehensibility. They can also be far more effective than more detailed indicators in putting real performance pressure on government agencies. This should not, however, blind us to the very considerable technical challenges involved in appropriately defining summary measures. Inappropriate selection of constituent variables, or poor technical design in respect to the aggregation of these variables, can produce summary measures which are quite misleading.

Linking Funding to Outcomes

In considering the role of performance measures in budgeting and funding decisions, it is important to explicitly consider certain key challenges which arise in linking outcome and/or outputs to funding. These have an important bearing on the choice of model of performance budgeting. We start by considering the outcomes/funding link, and after that look at the output/funding link.

In respect to outcomes, the impact of external factors on measured outcome variables has major implications for the type of linkage between outcomes and funding which it may be possible to create in a performance budgeting system. As noted above, most outcome indicators do not separate – or separate only to a limited degree – the outcomes achieved by government intervention from the impact of uncontrollable external factors. When the influence of external factors is extensive, it may be difficult or even impossible to predict the outcome which any particularly level of funding of a public service will produce. This is a major obstacle to attempt to create a *tight* link between funding and outcomes – for example, by basing a purchaser-provider system on outcomes. For this reason, performance bonus funding based on outcomes tends to be used only to provide small additional funding rewards (e.g. 5 percent additional funding), because if there was too strong a link between funding and unpredictable outcomes, the financial stability of the government agency concerned would be adversely affected. At the same time, outcome measures are crucial to program budgeting, and must constitute a key element in expenditure prioritization decisions.

This raises the question of outcomes targets and funding. The UK PSA system relied heavily on outcome indicators, over many of which the government had limited control. How, and to what

extent, can one link funding to targets which are only partially controllable? Clearly there has to be flexibility in the funding/target relationship under such circumstances. There are, however, many who take the view that flexibility is not enough and who dispute the wisdom of seeking to link budget decisions to *outcome* targets. In their view, target-setting in government should be confined mainly to *outputs* or other variables over which government has a high degree of control (e.g. activities).

Linking Funding to Outputs

The main focus in PB system such as formula funding and purchaser-provider is the creation of links between the quantity of output (i.e. volume of services provided) and the level of funding. For many outputs produced by government, there is a much stronger link between funding provided and outputs delivered (or deliverable) than is the case for outcomes. This is particularly true for *standardized outputs*, which are outputs where every client receives pretty much the same level of service, so that unit cost should the same. [15]

However, quite a few government services are not standardized. They are, rather, *heterogeneous outputs*. This means that the level of service provided to different clients, or in different cases, is deliberately varied so as to address differences in client conditions or circumstances. Police criminal investigations are a classic example – the amount of effort put in per case, even for the same types of case (e.g. murder investigations) varies enormously depending on the circumstances of the case. Even in school education, which is quite standardized for the great majority of students, heterogeneity is present when additional teaching and care activity is devoted to children suffering an intellectual or physical disability.

Substantial heterogeneity undermine the predictability of the relationship between the funding level and the outputs which the agency can be expected to produce, thus also influencing how tight a link it is possible to create between funding and outputs. This particularly affects the scope for applying formula funding and purchaser-provider systems, as discussed in section 12. This means that these forms of performance budgeting can only be applies selectively to the right types of services.

[15] At least if we leave aside cost difference arising from factors like geography.

There is one other type of service for which tight links between outputs and funding are problematic. This is *contingent capacity outputs*, of which a fire department is a good example. The fire department maintains capacity to provide at very short notice an output (firefighting) for which the demand is highly unpredictable. It would be unrealistic to seek to build a very close link between the number of fires attended by the fire service and the level of funding. Fire services cannot therefore be funded on a per-output basis, but must instead be funded in such a way as to deliver a certain level of *capacity* to fight fires.

The discussion to this point has focused on output *quantity.* There is also the question of potentially linking funding to output *quality.* It would be highly desirable to be able to include the output quality as well as quantity dimension in, for example, a purchaser-provider funding arrangement. One of the concerns about "perverse effects" raised by purchaser-provider systems is that, in funding only for output quantity, one creates incentives for agencies to cut costs by reducing quality. Including a quality component in funding could, in principle, resolve this problem. In practice, however, this is not easy, given the limits to our capacity to measure quality and the consequently highly imperfect nature of most quality measures. In general, the best hope for linking funding to output quality is through some element of performance bonus funding based on quality measures (similar to outcome bonuses) – in other words, by adding on to a system in which the main funding is based on output quantity a small additional element of quality-based funding.

Finally, there is the question of linking funding to activities – as suggested those who advocate "activity based budgeting". This will be discussed further in the session on formula funding. A key question which arises in basing funding on activities, however, is that of how far the activities which are being funding necessarily lead to results (outputs and outcomes) which the public cares about. Remember that while outputs are services, activities are "merely" work processes (see section 1). An organization which is internally-focused rather than externally focused on results may put excessive emphasis on activities for their own sake. To fund such an organization partly on the basis of internal support activities – such as meetings held or policy documents developed – could worsen this problem.

In conclusion, the tightness of the link which can be created between results and funding depends on the extent to which the underlying connection between funding provided and the results one can expect is clear-cut. If the relationship is an uncertain one – due to, for example, external factors or heterogeneity – then a tight link may be impossible. This is why, for example,

it is hard to use the purchaser-provider mechanism to fund organizations for outcomes. Conversely, program budgeting system will work even when there is considerable uncertainty in the relation between results and funding, precisely because program budgeting links funding to results only fairly loosely. This is why program budgeting is widely used as the basis for whole-of-government performance budgeting system, whereas mechanisms such as formula funding purchaser-provider can only work when applied to certain sectors or types of services (e.g. hospitals or schools).

The issues discussed in this section also have an important bearing on the role of indicators versus evaluation: namely, that the imperfections of performance measures mean that one cannot rely upon measures alone for the performance information to be used for expenditure prioritization.

Key Readings

Robinson, M (2007), "Results Information" in M. Robinson (ed.) *Performance Budgeting: Linking Funding and Results*.

5. Performance Measurement Systems

The performance measurement (PM) system refers to the mechanisms for data collection, the processing of data into indicators, the validation of those indicators (i.e. ensuring that the numbers are reliable and are not either manipulated or statistically questionable), and their presentation to users. A PM system is a subset of the broader monitoring and evaluation system.

In discussing the performance information requirements of performance budgeting, it needs to be borne in mind that performance budgeting is not – or should not be – an isolated reform. As emphasized earlier, it is part of a set of broader reforms, often referred to as "managing for results", which are designed to focus public management more on results delivered, and less on internal processes. These broader reforms include civil service reforms designed to increase the motivation and incentives of public employees; organizational restructuring to increase the focus on service delivery and improve coordination (e.g. creation of agencies and reduction of numbers of ministries); and institutional change to strengthen public accountability for performance. Action on these and a range of related fronts is necessary if the efficiency and effectiveness of public expenditure is to be tangibly improved. Improved performance information is fundamental to each of these elements of MFR reform because all need to be underpinned by better performance information. For example, a crucial ingredient in making the civil service more performance-oriented is the improvement of information on the activities and outputs of individuals, workgroups and agencies.

This implies that government-wide performance information strategy should be designed to meet not only the needs of performance budgeting, but of MFR processes more generally. It is not appropriate, for example, to attempt to develop a system of performance measurement aimed exclusively at budgeting applications, and entirely separate sets of measures used for civil service management, accountability or other purposes. The government-wide performance information system should be developed as an integrated whole. For example, if data collection relies on surveys, it will in general be much better to use one survey to collect a range of different information from a particular target group than to conduct a number of separate uncoordinated surveys.

Efficiency in the operation of the performance measurement system is particularly important because there are significant financial and human resource costs in the production of pertinent and reliable performance indicators.

The first stage in the production of performance indicators is data collection. The most readily available source of raw data for indicators is usually client service records such as hospital treatment records, school enrolment and attendance data, and data on regulatory approvals granted. A good PM system will need therefore, firstly, to improve the reliability of client service records and, secondly, to expand the type of data collected in clients service records so that it provides the basis for a broader set of indicators.

Client records provide some information about outcomes. For example, in schools, examination and graduations are recorded, and in hospitals deaths are recorded. However, in general, client service records provide information mainly about the volume of services (quantity of outputs) delivered to clients and the activities which these outputs comprise. A good PM system therefore needs to go beyond the constraints of client service records in order to construct more and better outcome and output quality measures. In this respect, the use of surveys (including, but not confined to, client satisfaction surveys) is the key element. Survey data can not only provide good measures of "outcome variables". It can also provide more timely measures, by capturing outcome variables which will only much later manifest themselves in demand for government services. For example, survey-based data on alcohol and tobacco consumption rates provides a much earlier warning of the extent of the problem which preventative health services need to address than do levels of alcohol and tobacco-related disease, which arises only after years of abuse. Surveys can pick up measure variables relevant to outcomes which take time to manifest themselves.

The second stage in the production of indicators is data processing – i.e. the transformation of raw data into performance indicators. A key question here is the choice of processing technology. There is a full range of options here, ranging from the most high-tech through to the relatively simple. At one end of the spectrum, performance measurement may be built into a large Integrated Financial Management Information System (see the section on program costing and accounting for more on IFMISs). The advantage of this – if the system works properly – is that the performance indicators are directly linked to the relevant financial data (e.g. program performance indicators are automatically linked to program expenditure data; output quantity measures to output costs etc).

At the other end of the spectrum, it is perfectly possible to manage a performance measurement system based on Excel spreadsheets and an Access data base. This is a simpler and lower-cost option. Some IT experts will suggest that incorporating performance measures in the IFMIS should be the preferred approach. However, the more complex and multi-faceted IFMISs are, the more chance they have of never working, or not working well. This is particularly true in developing countries. In many countries, the simpler option will work better.

The third stage in the production of indicators is indicator validation, which refers to the assurance of the reliability of the indicators. This is important to safeguard against errors and manipulation, as well as against methodological deficiencies which make the indicators unreliable. There are two aspects of indicator validation. The first concerns the validation of the capacity of agency systems to ensure indicator quality – for example, the adequacy of training and data entry checks.

Performance Measurement Protocols

The UK Royal Statistical Society proposed the idea of a Performance Measurement Protocol in its excellent 2003 report *Performance Indicators: The Good, the Bad and the Ugly*, suggesting that: "Before introducing performance monitoring in any public service, a PM protocol should be written. This is an orderly record not only of decisions made but also of the reasoning or calculations that led to those decisions. A PM protocol should cover objectives, design considerations and the definition of PIs, sampling versus complete enumeration, the information to be collected about context, the likely perverse behaviors or side-effects that might be induced as a reaction to the monitoring process, and also the practicalities of implementation. Procedures for data collection, analysis, presentation of uncertainty and adjustment for context, together with dissemination rules, should be explicitly defined and reflect good statistical practice."

The other aspect of indicator validation is the audit of indicators – that is, the assessment of the reliability of indicators by persons separate from, and independent of, the units responsible for their production. Audit may be internal (i.e. carried out by an audit unit within the relevant ministry), or external. External may mean that the audit is carried out by a government-wide body independent of the ministry which produces the indicators, or it may mean – going one step further – audit by a body independent of executive government (that is, a "supreme audit institution").

There are two broad approaches to the auditing of performance indicators. The first is the auditing of specific indicators. The other approach is a systems verification, which focuses on audit review of the performance measurement systems (including quality assurance processes) within each agency with the objective of forming an opinion as to whether these systems are such as to be likely to produce reliable performance indicators.

A quite separate issue – not to be confused with the *audit* of performance measures (which focuses on assessing their reliability) – is that of external review of the *appropriateness* of the indicators chosen by spending ministries. The choice of measures for performance budgeting purposes cannot be left to spending ministries alone, because the performance information has to serve the needs of central budget decision-makers. For example, if when the environment ministry is asked to come up with key performance indicators for its conservation program, it suggests the use of a measure of the number of policies which it develops, it should be firmly told that an *activity* indicator such as this is not satisfactory, and that it needs to develop *outcome* measures (such as independent estimates of the remaining populations of endangered species). The MoF and/or other central agencies must therefore provide guidance to spending ministries about the types of indicators they need to develop, and must subsequently reviewing the appropriateness of indicators suggested by ministries.

The question of whether the supreme audit institution should become involved in questioning the choice of indicator is more controversial. In some countries this happens (e.g. the US, where the Government Accountability Office frequently critiques indicators). By contrast, in British-like parliamentary systems, it is frequently asserted that to critique the choice of indicators would be to go too close to the critique of government policy, which is something which supreme audit institutions are traditionally barred from doing in those systems.

The final stage of the PM system is the presentation of the indicators. This needs to be done in such a manner as to make the indicators as readily understandable and usable as possible. Usability first and foremost depends on the selective of a relatively small number of indicators which are of greatest relevance to the user concerned. There is nothing which reduces the value of indicators more than presenting a user with hundreds of indicators most of which are of little interest to the user, but through which he or she is obliged to work in order to find what is of interest.

It is crucial in reporting performance indicators that they be placed into context, rather than presented simply as numbers in isolation. The means, in particular, the inclusion in a performance report not only of the indicators themselves, but of a narrative discussion of the trends which the indicators show and, in association with that, information on – and an assessment of extent of – the impact of relevant external factors. A "best practice" in external performance reporting – which not all countries will necessarily be able to implement, given financial considerations – is to create a multi-tiered system in which users can start at the broadest level of higher-level outcome indicators and then "drill down" to two or more lower levels of more detailed indicators, at the ministry and program level (many of which will be more focused on outputs, activities and input/resources).

Need for More Robust PM Systems in Africa

"A problem in African countries, and perhaps in some other Regions, is that although sector ministries collect a range of performance information, the quality of data is often poor. This is partly because the burden of data collection falls on overworked officials at the facility level, who must provide the data for other officials in district offices and the capital but who rarely receive any feedback on how the data are actually being used, if at all.

"This leads to another chicken-and-egg problem: Data are poor partly because they aren't being used; and they're not used partly because their quality is poor. In such countries there is too much data, not enough information. Thus, another lesson for the institutionalization of a government M&E system is the need to build reliable ministry data systems—to help provide the raw data on which M&E systems depend. An audit of data systems and a diagnosis of data capacities can be helpful in this situation. It would provide the starting point for any necessary rationalization of data collections or improvements in their quality. It would give the data the credibility necessary to be used."

Keith McKay (2007), *How to Build M&E Systems to Support Better Government*, World Bank Independent Evaluation Group.

Ensuring that Performance Information is Cost-Effective

The development of performance information systems is not simply a matter of developing the best and most comprehensive results and cost information possible. Rather, it is about a benefit/cost judgment. Performance information does not come free. It is costly in both financial and human capacity terms to design, build and then operate on a continuing basis the systems concerned. Careful judgments therefore need to be made about how far to go in respect to choices such as the number of performance measures to be developed, as well as

related questions such as the sophistication of program evaluation methodology and costing methodologies.

These choices face even the wealthiest countries. But they are particularly pressing for countries with more limited financial and skilled human resources. Such countries should be particularly selective and strategic in the development of performance measures. As discussed later, they should also, in many cases, make use of quite simple program evaluation methodologies. And, as mentioned above, they should not seek to develop efficiency measures which require relatively complex managerial accounting (e.g. allocation of indirect costs to outputs). The temptation of adopting what appear at the time to be cutting-edge practices — whether it be accrual accounting and budgeting at present, or purchaser-provider models ten years ago — should also be studiously avoided.

It is also a considerable challenge to build the capacity necessary to operate a good performance information system. For performance budgeting purposes this demands, in particular, great change in the skill set and competences of the MoF. Rather than being and exclusively economic/accounting body, the MoF must develop competence in policy analysis and in the development of performance information to support that policy analysis. Only in this way can it develop the capacity to advise executive government well about expenditure priority choices, in order to make effective performance budgeting possible.

Implementation strategy – including in relation to PM systems – is discussed in section 19. However, it is worth noting here that many developing countries are unrealistic in their expectations as to how long it should take to develop a comprehensive performance measurement system. Countries which are particularly advanced in this area – such as the UK and USA – took decades to get where they are today. It is a serious mistake for developing countries which are relatively new to performance measurement to set out to develop thousands of program indicators in very short time periods (e.g. 2-3 years). It is far better to start off modestly, focusing on a relatively small number of the most relevant indicators. Crucially, much attention should be paid to ensuring that collection, processing and validation systems are adequate in respect to that selective group of indicators. This is much more valuable than generating a much larger group of indicators which cannot be relied upon.

Key Readings

UK National Audit Office (2000), *Good Practice in Performance Reporting in Executive Agencies and Non-Departmental Public Bodies*, obtainable at http://www.nao.org.uk/publications/9900/good_practice_in_performance.aspx.

US Government Accounting Standards Board's (2005) *Government Service Efforts and Accomplishments Performance Reports: a Guide to Understanding,* obtainable at http://www.seagov.org/sea_gasb_project/sea_guide.pdf.

6. Evaluation and Performance Budgeting

Performance budgeting is often represented as being only about the use of *performance indicators* in the budget. This is wrong, because it overlooks the crucially important role of *evaluation*. It is for this reason that throughout this manual reference is made to the *performance information* base of performance budgeting, rather than to performance indicators alone.

Evaluation is the subject of another CLEAR course, and the general nature and methodologies of evaluation are therefore discussed here only very briefly. Further information can also be found in the readings listed at the end of this section. In general terms, however, evaluation is the formal assessment of programs, projects, organizations, or policies using systematic methodologies, with the intention of forming as objective an assessment as possible of their efficiency, effectiveness, design or management. One can distinguish between retrospective (*ex post*) evaluations, which are evaluations of programs etc which are already operating, and prospective (*ex ante*) evaluations, which are appraisals of possible new programs etc before they are implemented.

A Definition of Evaluation

"The systematic and objective assessment of an ongoing or completed project, program or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision–making process of both recipients and donors. Evaluation also refers to the process of determining the worth or significance of an activity, policy or program. An assessment, as systematic and objective as possible, of a planned, ongoing, or completed development intervention."

Keith McKay (2007), *How to Build M&E Systems to Support Better Government* (World Bank Independent Evaluation Group).

Evaluation can support the budget process by helping either to:

- Identify programs or components of programs which can potentially be cut: this means programs which are not cost-effective and which cannot readily be made cost-effective through policy design or management changes,
- Identify savings which can be made by improving the efficiency of service delivery.

Evaluating Program Effectiveness

The evaluation of program effectiveness has a particularly important role to play in those forms of performance budgeting which focus on the allocation of resources in the government-wide budget, of which program budgeting is the most important form. In such systems, the primary focus is upon making budgeting as performance-informed as possible. As we have seen, this means in particular that:

- Decisions about expenditure prioritization – where to allocate limited tax resources – are informed by good information on program effectiveness,

- Decisions about funding for specific ministries and agencies – and in particular decisions on their requests for additional resources – are informed by reliable information on how effectively the ministry or agency has used funding it has received in past budget.

Evaluation is crucial in this context because performance indicators are frequently insufficient in isolation to permit judgments on program or agency effectiveness. As we have seen, some program outcomes cannot be measured, or can be measured only very imperfectly, and many outcome indicators are heavily contaminated by external factors. Evaluation is very important as a means of making judgments about the likely impact of external factors on outcome variable. It can enable us to make some useful judgment about probable effectiveness of programs even when we lack any outcome measures (e.g. via program logic evaluation – see below). More generally, whether for budgetary or any other purpose, performance indicators alone rarely suffice (see box).

Role of Evaluation in a Balanced M&E System

"The most prevalent type of performance assessment practice … is indicator-based monitoring … Performance [indicator] monitoring, in and of itself, represents a relatively crude way to inform decision-making. In many cases, there is a need for a much more nuanced, in-depth understanding of the processes involved in particular programs or policies, which evaluations are much better equipped to provide. Given the specific strengths and weaknesses of monitoring and evaluation, which are potentially complementary to each other, the ideal approach is one that relies on an appropriate balance between the two types of activity."

Ariel Zaltsman (2006), *Experience with Institutionalizing Monitoring and Evaluation Systems In Five Latin American Countries*, World Bank Independent Evaluation Group.

Evaluation of effectiveness also an important role to play in systems of budget-linked performance targets liked the British Public Service Agreements system. As noted earlier, in such systems targets are set as an integral part of the budget process, with the aim of ensuring that the stringency of the targets is related to the extent of funding provided. After the event, scrutiny of actual performance against target is very important in this type of system, and it is here that evaluation can make an important contribution. For example, when the targets are outcome targets, a crucial part of the assessment of performance against target is consideration of whether, say, a failure to meet a target is the "fault" of the agency concerned, or whether it is instead due to the impact of unanticipated external factors. Evaluation can be very useful for this purpose. More generally, evaluation has a significant role to play in any performance management system which places considerable emphasis on the setting of performance targets, irrespective of how closely target-setting is linked to the budget.

There are two types of evaluation which are relevant to the assessment of program effectiveness:

- **Outcome evaluations** (also often referred to by evaluation specialists as "impact evaluations"): Outcome evaluations aim to directly measure the effectiveness of programs by taking outcome indicators and using sophisticated methods to assess the link between the program intervention and the measures changes in outcome indicators, in the process explicitly dealing with problems such as external factors. The methods employed include so-called "experimental" techniques (which entail the comparison of program beneficiaries with control or comparison groups at two or more points in time), and the use of regression analysis to separate the effect of external factors from other causal factors which impact on outcome indicators.

- **Evaluations of program logic**: these assess whether the program is designed in such a way as to make it likely that it will achieve its intended outcome. To evaluate program logic, the first step is to clarify exactly how the program is supposed to achieve its outcomes. Expressed in term of the "results chain", the key questions are: What intermediate outcomes is the program expected to deliver? How is it that those intermediate outcomes are expected to generate, or contribute to, the program's intended higher-level outcomes? Once the program logic is clarified, the next step is to ask whether it is reasonable to assume that the program will achieve its intended outcomes. For example, given what we know about relevant economic theory, is it reasonable to assume that a specific industry policy will deliver its intended outcomes?

Efficiency Review

The use of evaluation to identify savings from improved efficiency of service delivery in relevant to any form of budgeting. In this context, the form of evaluation we are talking about is **efficiency review.** To the extent that efficiency reviews go beyond recommending steps to improve efficiency to providing quantified estimates of potential savings, they are of potentially great value to ministries of finance.

Efficiency review has a particular relevance to forms of performance budgeting such as formula funding and purchaser-provider systems, in which funding is based on the unit costs at which the funding authority believe the service delivery agency should be able to produce outputs. Specifically, the findings of efficiency review can inform that funding authority about what unit cost to use for funding purposes. For example, suppose that the government has been funding the health ministry for vaccinations at the ministry's present unit cost $20 per head, but that an efficiency review indicates that this current unit cost is greater than it should be and could be reduced by 10 percent over two years. Under these circumstances, it would be reasonable for the funding authority to plan to gradually reduce the amount it pays to $18.

Although efficiency reviews are useful for this type of performance budgeting, evaluations of effectiveness clearly are not. This is because these types of performance budgeting focus are directed to improving efficiency rather than improving the allocation of resources.

Making Evaluation Relevant to the Budget

Informing the budget process is not by any means the only role of evaluation. Broadly spending, one can distinguish valuation for budgetary purposes from evaluation for policy/management improvement purposes. The latter type of evaluation aims to help institutions improve policy design – that is, help institutions to change the nature of the services they deliver to the community so as to make them more effective in achieving their intended outcomes – without focusing on identifying programs which can be cut. Or it can aim to help institutions improve processes and management so as to make the delivery of services more efficient, but without a focus on quantifying budgetary savings which can be made as a result of such efficiency improvements.

A good performance budgeting system therefore requires the conduct of selected evaluations *specifically intended to inform the budget process* – that is, designed to give budget decision-makers better information upon which to base budget decisions. Expressed differently, it cannot be assumed that, simply because a government conducts substantial evaluation work, that this evaluation work will necessary meet the needs of performance budgeting.

This has implications, firstly, for the selection of topics to be evaluation. If evaluation is going to make its maximum potential contribution to resource allocation, the programs (or elements of programs) and topics chosen for evaluation should be those which appear prima facie likely to yield budgetary savings.

Secondly, budget-linked evaluation needs to deliver its findings quickly and at the right time to be taken into account in budgetary decisions. It is important here to bear in mind that evaluations can be conducted in great depth, if desired, making extensive use of surveys, interview and other data gathering techniques, in combination with quite sophisticated analytic techniques. Such in-depth evaluations may generate more reliable results, but it tends to take considerable time, which may mean that it is not well geared to serving the needs of budget decision-makers.

Because of the important of timely information, considerable emphasis is being placed these days on *rapid evaluations* for budgetary purposes. Rapid evaluations are evaluations carried out in quite short time frames (e.g. 3-6 months). Rapid evaluations tend to focus on the evaluation of program logic rather than outcome evaluation. As indicated above, evaluations of program logic aim to assess how plausible it is that the program would generate its intended outcomes. They are in principle inferior to outcome evaluations. However, outcome evaluations tend to take considerable more time to produce and to cost considerably more. The speed and low-cost character of rapid evaluation makes it very attractive in practical times, particularly – but by no means exclusively – in low income countries.

Evaluation can make its greatest contribution when linked to systematic spending review (SR), which is discussed in Section 15. SR refers to the systematic scrutiny of ongoing program expenditure to identify options for cuts. The MoF will in general manage and conduct spending review, providing advice on potential cuts to the political leadership. In the process, the MoF can make best use of evaluation by commissioning rapid evaluations focused on programs which it or the political leaderships feels may be of doubtful effectiveness.

Evaluation and the Budget: The Chilean Example

The best developed example of budget-linked evaluation linked to spending review is that which has been progressively developed in Chile from the late 1990s. In Chile, performance budgeting is part of a broader performance management system, known as the "system of evaluation and management control". This system aims to improve the effectiveness of policy-making and management throughout central government, to create performance incentives for civil servants, and to make the budget results-oriented. The degree of use which Chile makes of systematic evaluation as a basic tool of its performance management system is quite

exceptional by contemporary international standards, at least outside Latin America. Evaluations are managed by the MoF, and the topics for evaluation are centrally determined. There are three different types of *ex post* evaluation in the Chilean system:

- Outcome (Impact) Evaluations: for example, an evaluation of whether a labor market program had achieved its stated objective of bringing the long-term unemployed reintegration back into the workforce, carried out mainly via a "longitudinal" analysis of the track record of participating long-term employed persons to ascertain the rate at which they obtained and retained work, and how this compared with other long-term employed persons who did not participate in the program.

- Evaluations of Government Programs: the main element of which is the evaluation of program logic (see above).

- Institutional Evaluations – going under the potentially misleading name of "comprehensive expenditure evaluations" (Evaluaciones Comprehensivas de Gasto), these evaluations look at specific institutions or sectoral groups of institutions. They examine a range of issues including the consistency of institutional and sectoral objectives, organizational structures, production and management processes, resource use and service-delivery performance.

The coverage of *ex post* evaluation has progressively increased over time, with 33 evaluations in 2009 and 39 in 2010. The programs and organizations to be evaluated are selected by the MoF in consultation with the Congress. Evaluations are then carried out by external evaluators (consultants or research institutions) contracted by the MoF, which provide clear terms of reference and methodological guidelines to the evaluators. All final evaluation reports are made available to the Congress and public, and their summaries are included in the budget information papers in the form of "Executive Minutes". The MoF and the relevant ministry discuss the recommendations of the evaluations and agree on the actions which should be taken in response to evaluation recommendations. This then becomes the subject of a formal agreement, the implementation of which is the monitored in subsequent years by MoF.

In addition to *ex post* evaluation, Chile has recently extended its evaluation system to *ex ante* analysis of new spending proposals. The MoF requires institutions to present all new spending proposals in a standard format designed, amongst other things, to make absolutely explicit their intervention logic. "New spending proposals" refers not only to proposed new programs, but also to significant discretionary expansions of existing programs. Associated with this, MoF introduced in 2009 a new formal mechanism of *ex ante* evaluation of new spending proposals. More recently, it has added to this a technical assistance service to entities under which MoF provides advice on how to develop and present good quality new spending proposals.

Internal versus External Evaluations

Evaluation for budgetary purposes is only one part of a well-developed evaluation system. It is, in particular, appropriate that all ministries and agencies be encouraged to carry out systematic programs of internal evaluations as a management improvement tool. Many countries have developed government-wide evaluation policies to achieve exactly this. [16] The focus of such spending ministry-led evaluations will tend, in general, to be more operational than that of evaluations initiated by the Ministry of Finance or other central government agencies to inform budget decisions. It is inappropriate for the MoF and political leadership to rely on evaluations led by ministries themselves as the information source for budget decisions. Such evaluations not only have a more operational focus, but they will not necessarily be focused on the programs which are the most likely candidates for budget cuts. Moreover, the knowledge that the results of the internal evaluations will be used by the center to cut programs may lead ministries to manipulate evaluation conclusions to make sure that they are not too damaging. In a well-developed evaluation system, the program of specifically budget-linked evaluations should therefore be only the tip of the iceberg of a much broader evaluation effort.

In conclusion, evaluations is critically important as part of the performance information base for performance budgeting. However, to be useful, evaluation has to be of the right type – which means both that it is timely, and that it is geared to answer the questions of most relevance to budget decision-making.

Key Readings

Guzmán, Marcela (2007), "The Chilean Experience" in M. Robinson (ed) *Performance Budgeting.*

HM Treasury (1997), *Appraisal and Evaluation in Central Government,* London: The Stationary Office, available at http://www.hm-treasury.gov.uk/d/green_book_complete.pdf.

McKay, Keith (2007), *How to Build M&E Systems to Support Better Government*, Washington: World Bank, available at http://www.worldbank.org/oed/ecd/better_government.html.

OECD (1998), *Best Practice Guidelines for Evaluation*, PUMA Policy Brief No. 5, Paris: OECD, available at http://www.oecd.org/dataoecd/11/56/1902965.pdf.

[16] For a case study of the Mexican system, see Manuel Fernando Castro et al (2009), *Mexico's M&E System: Scaling Up from the Sectoral to the National Level*, ECD Working Paper Series No. 20, Washington, World Bank (obtainable at www.worldbank.org/ieg/ecd).

World Bank (2004), *Monitoring and Evaluation: Some Tools, Methods and Approaches*, Washington: The World Bank, available at http://lnweb90.worldbank.org/oed/oeddoclib.nsf/a7a8a58cc87a6e2885256f1900755ae2/a5efbb5d776b67d285256b1e0079c9a3/$FILE/MandE_tools_methods_approaches.pdf.

7. Performance Auditing

Defining performance auditing is not straightforward, because it is not possible to distinguish performance auditing *methodologically* from evaluation. In practice, the term is usually used to refer to evaluations which are carried out by supreme audit institutions (SAIs). SAI is the generic term used internationally for bodies such as the UK National Audit Office, the US Government Accountability Office, and the French *Cour des Comptes*, which are independent (to varying degrees) of executive government and designed to hold it to account. However, within the audit profession, the term "performance auditing" is sometimes also used to refer to evaluation carried out by *internal* auditors – that is, by auditors who are part of the executive government bodies which they audit. It may therefore be that the best definition of performance audit is that it is performance-related evaluation carried out by individuals or institutions with a primary professional background and responsibility for financial audit.

Because SAIs are not part of executive government, performance auditing carried out by SAIs are by definition not intended to serve (other than incidentally) as instruments to aid executive budget preparation. Generally speaking, such SAI performance audits are better seen as instruments to assist the parliament in more effectively carrying out its performance accountability role. This distinguishes them fundamentally from executive government evaluations systems, including the type of budget-linked evaluation which ministries of finance should be building as part of the performance information framework of performance budgeting (as discussed in the evaluation section). However, in those countries where parliament has significant independent power over the allocation of resources in the budget, SAI performance audits can also have a more direct influence on the budget and be seen as an integral part of the performance budgeting system. The section on parliament and performance budgeting discusses these different parliamentary budget roles.

It is useful to distinguish two different types of performance audits, the first of which is the *substantive performance audit.* Such audits aim to form explicit judgments on the effectiveness and/or efficiency of programs, government agencies or functions. The other is the *system performance audit*, which focuses instead on assessing the extent to which specific categories of management system or practice are conducive of effectiveness and efficiency. Substantive performance audit tends to be more common, and more wide-ranging, in political systems characterized by the separation of powers between the parliament and executive, and consequently be considerable parliamentary independence vis-à-vis executive government. By contrast in parliamentary systems where government ministers tends to control parliament, there is sometimes a tradition that the SAI should not challenge government policy, which

(depending on how “policy” is interpreted) can have the effect of narrowing significantly the scope of substantive performance audit.

In a number of countries, a type of executive performance auditing carried out by executive government bodies which are separate from the ministries which they audit. For example, the Indonesian Financial and Development Supervisory Board – a body like an SAI but part of external government (there is a separate SAI which reports to parliament) – carries out performance audits. And in France, the introduction of the new performance budgeting system in 2005 was accompanied by the development of a form of executive systems audit carried out by an Interministerial Committee for Program Audit, and focused on issues crucial to the evolution of the new systems.

Key Reading

Shand, D. (2007), “Performance Auditing and Performance Budgeting” in Robinson (ed.) *Performance Budgeting.*

8. Program Budgeting Fundamentals

Program budgeting is the planning, authorization and execution of expenditure in terms of *programs*. Programs group together expenditure on specific public policy purposes, such as environmental conservation or higher education. The classification of expenditure in terms of programs turns the budget into an instrument for explicit choices about expenditure priorities such as how much to spend on preventative health vs. treatment health; how much on tertiary education vs. primary education; and how much on strengthening the army vs. promoting agriculture.

Improving expenditure prioritization is therefore the primary objective of program budgeting (although, as discussed below, program budgeting also acts by increasing the pressure on ministries to perform). Expenditure prioritization refers to the allocation of funds to the sectors and programs which are most effective in meeting social needs (see Section 1). In the public sector, decisions about the allocation of resources are to a large degree be made by planning – a process whereby either the government as a whole, or individual spending ministries, decide what types of goods and services will be provided to the community, and to whom. This contrasts with the market economy, where the allocation of resources between competing products and services consumer is largely decided by customer demand, without the need for conscious planning. Program budgeting is an instrument for integrating planning and budgeting.

Program budgets are very different from traditional budgets, in which funds are allocated to each spending ministry not by objective but instead primarily by “line item”. Line-items are allocations of funds to types of inputs such as salaries, supplies and capital expenditure. The problem with a budget based on line items is that it indicates next to nothing about the policy objectives of the expenditure concerned. A line-item budget is therefore essentially useless as an instrument for expenditure prioritization. It was precisely to overcome this weakness that the concept of program budgeting was originally developed. Program budgeting also differs from traditional budgeting in that it calls for the substantial reduction of line-item controls over how spending ministries use their budgets. This is because program budgeting – and performance budgeting more generally – call for greater freedom at the ministry level in the choice of the inputs used to deliver services in return for greater accountability for the results which ministries deliver to the community. This does not, however, mean that budget allocations to line items entirely disappear under program budgeting (see Section 9).

Defining Programs

To facilitate improved expenditure prioritization, programs (and their constituent sub-programs – see below) need to be defined in such a way as to capture the choices about spending priorities which are made at the government-wide level (i.e. by presidents, cabinets, ministers and ministries of finance), and by the spending ministries themselves.

To capture such choices, programs are, first and foremost, categories of expenditure directed at achieving a common outcome. For example, a nature conservation program covers expenditure on a range of interventions such as the enforcement of laws banning the hunting of native species, marketing campaigns designed to raise public awareness of the importance of protecting the natural environment, actions to prevent the destruction of natural environments which endangered species depend upon, and the "culling" of feral species (e.g. cats, introduced fish etc) which may threaten native species. What all these interventions have in common is that they all aim to preserve native fauna and flora – which is an outcome.

While programs are defined in large measure in terms of the common outcome which they aim to achieve, in many cases they have other defining characteristics which also relate to the choices about expenditure priorities which governments and spending ministries wish to make. Budget decision-makers are, for example, often also very much concerned with *who* benefits from expenditure. For this reason, programs are some cases defined in terms both of their intended outcome *and* their target client group or region. For example, education ministries typically have separate programs for primary and secondary education. These programs have the same intended outcome – educated and socialized young people – but they target different client groups (i.e. young people in different age ranges). Similarly, in health treatment services, the government may wish to make explicit decisions about the regional allocation of health budget resources – the result of which may be a program structure in which there is a hospital program which is comprised of regional sub-programs.

Similarly, programs are sometimes defined in terms of both their intended outcome *and* the mode of intervention used for achieving this outcome. For example, the intended outcome of a "preventative health" program is reduced death and disability from disease and accidents, but a defining characteristic of this program is that it aims to achieve this outcome by preventative means.

Programs are linked to the "results chain" not only via the outcomes concept, but also via the concept of *outputs*. As explained in Section 3, outputs are goods or services – the "products" – which a ministry or other government organization delivers to external parties. In stating that a

nature conservation program is comprised of a range of different "interventions" aimed at protecting native flora and fauna, what we are saying is that the program is a group of different *outputs* which share a common intended outcome (as perhaps other common characteristics such as a common client group or mode of intervention). To take another example, a "crop industries" program groups together a range of outputs such as extension services, fertilizer subsidies and marketing support all of which have the objective of boosting the crops industry. And a preventative health program might include outputs such as sanitation promotion publicity campaigns, the placement of notices warning people against swimming or washing in lakes or rivers with waterborne diseases, and the promotion of exercise and fitness in the community. Listing the most important outputs (types of services) which fall under each program is an important part of developing a good quality program classification.

In this context, it is important not to make the mistake of confusing outputs with activities, and of viewing programs as being comprised of constituent activities rather than constituent outputs. As explained in Section 3, activities are the work processes which are used to produce outputs. For example, in health treatment, it is the treatments received by patients which are the outputs, and the activities include nursing, surgery, cleaning and the maintenance of hospital records. Programs are defined in terms of what matters most to the community – the services (outputs) received by clients, and the outcomes which those services generate. They are not defined in terms of activities and inputs, which are more of relevance for internal management purposes within ministries.

In summary, programs are first and foremost groups of outputs with shared outcomes, although they may also have other defining characteristics. Only when defined in this way are programs a useful tool for expenditure prioritization. There are, as discussed below, limited exceptions to the principle of outcome/output based programs, and these exceptions are driven by purely pragmatic considerations which have nothing to do with the basic objectives of program budgeting.

Performance Pressure and Program Budgeting

Although the primary objective of program budgeting is improved expenditure prioritization, this is not the only way in which program budgeting improves public sector performance. A well-designed program budgeting system also improves the effectiveness and efficiency of expenditure by putting additional pressure to perform on ministries and other budget-dependent government organizations. Under program budgeting, each ministry's budget

request to the MoF has to be accompanied by information on the performance of the ministry's programs (indicators and evaluation findings), as well as – where appropriate – performance targets for the future. The knowledge that its budget request will be considered by the government and, perhaps, the parliament in the light of its performance should then put increased pressure on the ministry to improve its performance.

Program Performance Information

Program budgeting clearly requires not only *budgeting* in terms of programs, but also the systematic use of program performance information. Only through the development of good program performance information does it become possible to compare the budgetary cost of each program with the results which the program delivers to the community.

This means, firstly, that program budgeting requires the systematic development of performance indicators for each program (or sub-program). Consistent with the expenditure prioritization objective, these program performance indicators should as far as possible measure program outcomes and outputs. The following provides an example of the types of performance indicators which need to be developed for an effective program budgeting system.

Health Treatment Program	
Outcome indicator	• Weighted average survival rate after treatment for selected major life-threatening conditions
Output quantity indicator	• Number of patients treated
Quality indicator	• Patient satisfaction rate, • Readmission rate after treatment, • Weighted average waiting time for selected major surgical procedures.
Efficiency indicator	• Weighted average treatme diagnoses.

One of the common deficiencies of program budgeting systems in many countries is that program performance indicators are dominated by indicators of activities and inputs, such as numbers of policies developed, numbers of meetings held and numbers of positions filled. These types of indicators are very useful for internal management purposes, but they are not the type of indicator needed to serve the objectives of program budgeting. The reason for the dominance of this type of indicator is usually that they are readily available, whereas good indicators of, say, outcomes and output quality are often not easy to develop. But using such indicators simply because they are available is like the old story of the man who, having lost his

watch in a dark corner of a street at night, decided to concentrate his search under the street lamp because he could see clearly there.

Program evaluation is at least as important an element of program performance information as are program performance indicators. For the purposes of expenditure prioritization, the most important performance consideration is the *effectiveness* of programs and their constituent outputs. As noted in Section 6, performance indicators are rarely sufficient in themselves to permit effectiveness of public expenditure to be assessed. Outcome indicators typically suffer from many imperfections, one of the most important of which is that they rarely adjust, or adjust fully, for the impact of "external factors". Where this is the case, only evaluation can help to distinguish the impact of external factors from that of the government's own efforts.

Unfortunately, it is all too common internationally for governments to neglect the performance information side of program budgeting and to believe that, merely by developing a program classification of expenditure, they have implemented program budgeting. However, it cannot be emphasized too strongly that merely knowing how much programs cost is of little use without information about the benefits which programs deliver. Moreover, performance indicators alone are insufficient for this purpose – particularly if they are mainly indicators of activities and inputs.

Program Titles and Objectives

Each program is defined by its title, overarching objective and the key outputs [17] which it includes. The program title should be short and informative. It should make as clear as possible to the political leadership, parliament and the public what the program is. Depending on the program, this might be done best by referring to the type of outputs, client, or objective of the program. Examples of good program titles are: "nature conservation", "crime prevention", "adult literacy", and "curative health".

Clearly defining the overarching program objective is important not only for clarity in program definitions, but also to provide a framework for the derivation of program performance indicators and targets. Program objectives should make clear the outcome which the program as a whole seeks to achieve, as well as (if relevant) other defining characteristics of the program such as the target client group or mode of intervention. [18] The program objective must be a

[17] Or, exceptionally, support services in the case of "administration programs (see Section 11).

[18] Because the outcome is the core of the program objective, one might ask whether it would make things clearer if spending ministries were told to develop statements of "program outcomes" rather than "program objectives". Some countries do exactly this. There is, however, one very good reason for referring to program "objectives'',

statement of the objectives which are shared by all of the different services grouped together within the program. Subordinate, more "operational" objectives may [19] also be defined for each program, and these will in general relate to some but not all of the services covered by the program. The discipline involved in requiring spending ministries to define an *overall* objective for each program helps to ensure that programs are – other than in the case of certain defined exceptions – appropriately defined with reference to outputs and their common outcome.

Program objectives should be explicitly and briefly – ideally in a sentence – defined in the budget documents. There are many cases internationally of program objectives which are not well defined. It is, for example, not unusual to find program objectives which focus entirely on the output (service) which the program delivers to the public, or on program activities/processes, with no reference to the intended outcomes (see below). Sometimes, program objectives are formulated in a wordy and unclear way, perhaps with excessive use of bureaucratic language.

The test of a common outcome which can be expressed in an overall program objective is an important test of the integrity of the program definition. For example, if the army health service has over time expanded its role from one of providing health services to soldiers to one of direct health service delivery to the public – in competition with the health ministry – it would not make sense to include the army health service in an "army" program, the outcome of which should be defined in terms of national security.

Right and Wrong Ways of Defining Program Objectives

Examples of well-formulated program objectives are:

- "The conservation of biological diversity in healthy ecosystems" (Nature Conservation program),
- "Maintenance of territorial integrity and national independence" (Armed Forces program),
- "Increased foreign investment leading to technology transfer and a stronger economy" (Investment Facilitation program),
- "Reduced crime and greater security of persons and property" (Crime Prevention program).

rather than to program outcomes. This is that there are certain exceptional programs which do not deliver services directly to the public (see below), and which do have aim to achieve specific outcomes, but rather to support other programs which are outcome-focused. Such programs still have objectives, but these objectives are usually not outcomes.

[19] In fact, in a fully developed program budgeting system, such more detailed objectives should definitely be developed. However, this need not be done at the outset in the implementation of such a system, whereas overarching objectives should be defined simultaneously with programs themselves.

Examples of the wrong approach to defining program objectives are:

- "Provision of medical assistance to persons in an emergency" – refers only to outputs.
- "Manage the development, implementation, evaluation and maintenance of national policy, programs and systems for general education and quality assurance" – refers only to activities.

Program Statements

To achieve its aim of improved expenditure prioritization, an effective program budgeting system must bring information on the performance of programs – that is, on their success in achieving their intended results – together with information on their cost. Being able to see the results achieved by programs alongside their cost helps budget decision-makers to make better judgments about whether programs should be cut, expanded, or maintained.

Parliament and the public should be kept informed via program statements presented with the budget documents which accompany the annual budget legislation. Program statements should include the following information for each program:

- Title and objectives,
- List of the main outputs (services) which comprise the program,
- A brief narrative outline of program strategy, challenges and key new initiatives,
- Key program performance indicators,
- Program performance targets, if applicable,
- Program expenditure estimates, preferably with medium-term projections,
- A breakdown of program expenditure by broad categories of economic classification (staff, capital etc), for information purposes.

The specific content of the program statement in respect to a number of these elements is discussed in more detail in subsequent sections.

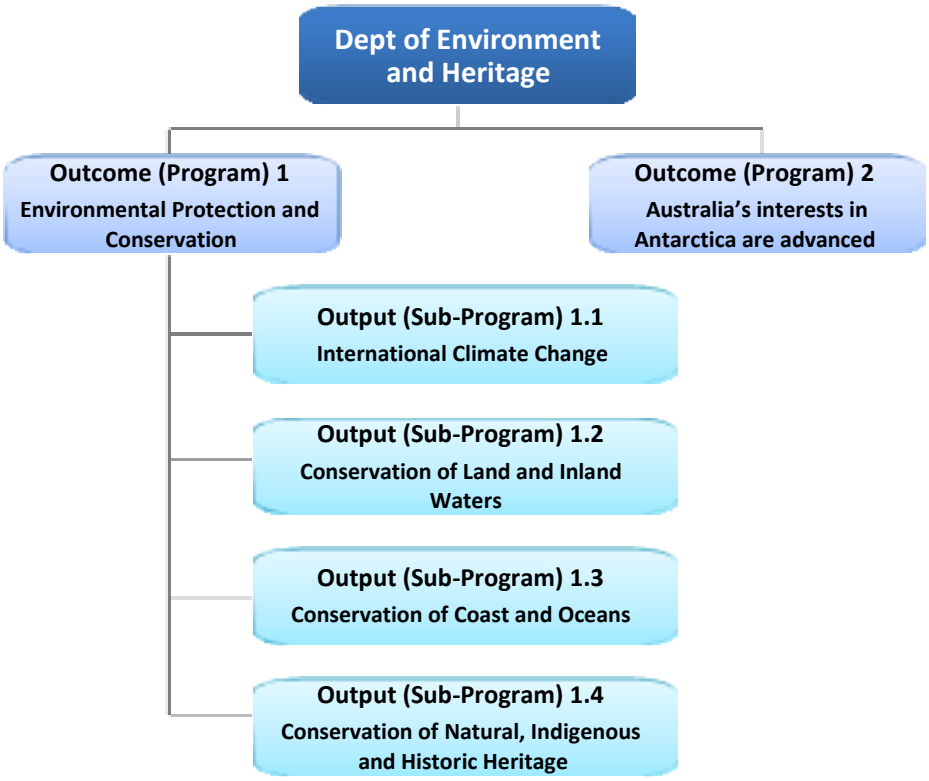
One way of presenting this material is for each ministry to prepare a document containing all of its program statements to be made available to parliament as an annex to the budget documents. The MoF should determine a standardized format for these documents.

The Program Hierarchy

Programs are typically decomposed into constituent "sub-programs", and sometimes even into lower levels (sub-sub-programs etc) – although it is suggested in later sections that it is better

to keep the number of levels limited. An example of this type of "program hierarchy" is shown in the box below. Note that the terminology employed for these elements of the program hierarchy varies greatly around the world (as does the terminology used for programs). The following is an Australian example from a few years back showing how programs (in this case called "outcomes") are broken into sub-programs (called "outputs") which were further decomposed into sub-sub-programs (called "sub-outputs").

From Program to Sub-program: an Australian Example



Subprograms and other lower-level elements of the program hierarchy are used mainly for internal management within the ministry or agency concerned. For the most part, they represent a level of disaggregation of expenditure which is too detailed for the central decision-makers to concern themselves with during the budget preparation process.

Number and size of programs

Programs are the level at which central decisions about expenditure priorities will generally be made. This has two important implications when formulating the program structures for ministries:

- Creating one big program covering all or most of a ministry's expenditure is generally a mistake. For example, it is better to have a number of programs such as primary education, secondary education and tertiary education in an education ministry than to have a single enormous education program. And rather than a single agriculture program, an agriculture ministry should have a number of programs, perhaps along industry lines (i.e. separate programs for crop industries, livestock industries, and fisheries). A program classification based on giant programs will be too coarse to serve central decision makers in making strategic spending reallocations such as, for example, shifting money from tertiary education to primary education or from treatment health to preventative health. Expressed differently, programs should capture each of the distinct key aspects of the role of each ministry so as to permit central decisions about the broad lines of the ministry's service delivery focus. It is in no way inconsistent with this to note that, for small ministries with narrowly-focused missions, a structure with a single program may be appropriate.
- Too many small programs should be avoided. Because central decision makers need to concentrate primarily on the "big picture" of expenditure prioritization across government, a proliferation of very small programs runs the risk of unnecessarily complicating the central budget preparation process.

A consistent approach to the size and number of programs needs to be applied across government. It would, for example, be inappropriate if the health ministry decided it would have a single program, while the justice ministry decided to have twenty. Guidelines from the MoF on this are essential (see further below in this section). Inevitably, such guidelines will need to have recourse to rules which are arbitrary but necessary – in particular, a limit on the number of programs any ministry have (perhaps something between 5 and 8, but this depends on how many ministries the government has), and a minimum size for programs (e.g. not less than a certain value, or alternatively no less than a certain percentage of the ministry budget).

Developing Programs

The proper definition of programs and the program hierarchy is a task which must be got right if program budgeting is to work properly. This means that the MoF cannot simply leave it to ministries to determine what their programs will be. The MoF must, rather, issue clear guidelines covering issues such as the formulation of program objectives, program size and number, and whether there will be a single administration program in each ministry (see later

sections). Only in this way will the interests of central budget decision-makers in having programs which help them translate government priorities into budget numbers be respected.

On the other hand, it would be wrong for the MoF to dictate to each ministry, without consultation, what its programs will be. Spending ministries know their business better than the MoF does, and it would be a mistake not to benefit from spending ministry expertise. Moreover, the program structure – particularly at the lower levels of the program hierarchy – is intended to serve as an internal budgeting and management tool for the spending ministries themselves. This will not happen if the spending ministries feel no sense of ownership of the program structure.

In summary, the development of the program structure must be a collaborative endeavour based on a partnership between the relevant spending ministries and the MoF.

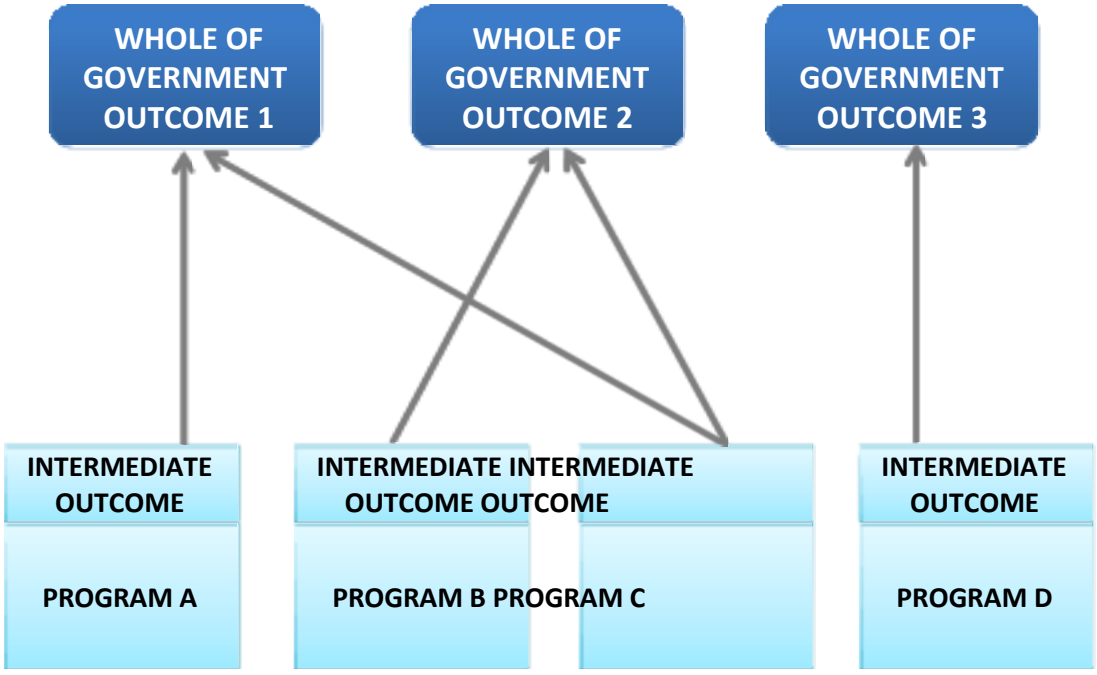
Integrating Strategic Plans with the Program Budget

Program objectives need to be explicitly linked to the objectives formulated in the organization's strategic plan and in any government-wide strategic or national plans. More generally, planning and programming should be seen as part of an integrated cycle.

A well-formulated government-wide strategy will define a small number of high-level outcomes upon which the government is focused. The clear specification of program objectives then provides the natural means of linking programs to the government-wide strategy. This is because program objectives are generally "intermediate" outcomes through which the high-level whole-of-government outcomes are achieved. For example, the outcome of a Criminal Policing program ("reduced crime") would link it directly to a whole-of-government outcome such as "a safe and secure society". Similarly, the outcome of the High School Education Program ("young people educated and socialized to take their place in the workforce or higher education") would link it directly to whole-of-government outcomes such as "rapid and sustainable growth of GDP".

This type of link between the programs and national strategic outcome can be represented graphically as illustrated the figure below.

Linking Programs and National Strategic Outcomes



Strategy should also be linked with the program budget at the entity level. Canada, France and the United States are amongst the leading countries which have fully integrated entity strategic plans with the budget by including within the budget documents entity performance plans which are:

- Based on the program structure of the budget: in other words, performance objectives, plans and targets are outlined for each of the entities programs,
- Covers all entity spending, not just "new" spending,
- Include only entity spending initiatives for which the government has approved funding in the budget.

These three countries also require that entities present annual performance reports which report on success in achieving the objectives set out in the performance plans.

In the US, this approach has been mandatory since Congress passed in 1993 the Government Performance and Results Act (GPRA). The Performance Plans which GPRA requires are not full entity strategic plans, but GPRA stipulates that they must be consistent with the entity strategic plan.

The structure of the French *Projets Annuels de Performance* (annual performance plans) is slightly different in that it is based not on entities but on "missions", which are groups of programs which in some cases spread over more than one entity. Apart from that, the French performance plan structure is very similar to other countries. It is broken into sections for each budgetary program, with presentations of program objectives, strategy, indicators, and expenditure.

Elements of this broad approach to integrating entity strategic plans with the budget are to be found in many other countries. In New Zealand, for example, the government's budget documents include entity Statements of Intent which are essentially strategic plans structured around entity-specific outcomes (equivalent to programs). In South Africa – one of the many developing countries which have adopted a similar approach – regulations also stipulate that each entity's "strategic plan must … be consistent with the institution's published medium term expenditure estimates".

Key Readings

Robinson, M. and H. van Eden (2007), "Program Classification", in Robinson (ed.) *Performance Budgeting.*

Kraan, D.-J. (2008), "Programme Budgeting in OECD Countries", obtainable at http://siteresources.worldbank.org/INTPRS1/Resources/383606-1201883571938/ProgrammeBudgeting_OECD.pdf .

9. Program Appropriation and Expenditure Control

Program budgeting involves not only planning, but also approving and executing the budget in programmatic terms. In this section we focus on the nature of the budget authorization under program budgeting and the degree of flexibility in respect to that authorization during budget execution.

Under most program budgeting systems, the parliament votes the budget allocations in programmatic terms, and this programmatic allocation constrains executive government in the execution of the budget. The parliamentary authorization is in most countries only at the program level. Only in a minority of countries does parliamentary appropriation of the budget take place at sub-program level, which tends generally to be thought of as areas of executive government discretion. [20]

In considering the nature of the enacted budget, however, we should have regard not only to the form in which parliament approves the budget, but also the additional "control totals" which the minister of finance or other executive bodies might impose. For example, in a few countries, even though the parliament only approves the budget at the program level, the minister of finance approves and controls the allocation of programs between sub-programs.

A key question under any program budgeting system is therefore that of how detailed a control to exert at the center (whether by parliament or central executive bodies). One approach to this question is to recognize that the information available to central decision-makers is never going to be sufficiently strong to permit them to determine all of the detail of resource allocation across the whole of government. Broadly speaking, the center is better of concentrating on the "big picture" of resource allocation, while leaving considerable discretion to spending ministries on the detail. From this perspective, allocation at the sub-program level may be better left to ministries themselves to determine.

Central approval of the budget in terms of programs raises the question of *transfers* (sometimes known as *virements*) of expenditure between programs. This refers to transfers of money between categories of funding approval – in this case, between programs. In many program budgeting systems, parliament authorizes executive government to make limited

[20] There are exceptions to these generalizations. One is the "global appropriations" system of countries like Australia and Canada, under which the programmatic allocation is merely indicative and can be changed by executive government at its discretion.

modifications to the approved program allocation of funds – for example, to move up to 5 percent of the funding of any program to other programs without having to come back to parliament to modify the budget. This type of transfer arrangement is intended to create a little flexibility to respond to unanticipated developments which might occur during the year.

The other key issue which arises under program budgeting is the reform of the traditional budget appropriation framework. As mentioned earlier, traditional budgeting is based primarily on "line-item" approvals. This means legal allocations of funds to inputs – known formally as "economic classifications" of expenditure – such as salaries, supplies and capital expenditure. In traditional systems, the line-item allocation is sometimes very detailed. For example, there may be separate fund allocations to items such as office stationary supplies, books, internal training, external training etc. This means that spending ministries are then unable -- at least without central approval – to shift money between these categories during the year. In practice, some flexibility is created by systems which permit some transfers between line-item categories. However, not only is the scale of such transfers typically greatly limited, but the process of reviewing and determining transfer applications can involve a great deal of unproductive work for the MoF.

Performance budgeting is opposed to his type of *detailed* control over the line-item composition of expenditure. The performance budgeting focus is on the results delivered by agencies, not on the way on which they deliver those results. Performance budgeting calls for ministries, and managers within those ministries, to be given greater freedom to choose the best mix of inputs with which to efficiently and effectively deliver services. The corollary of this is increased accountability for the results delivered. [21]

For this reason, performance budgeting has typically been accompanied by extensive reduction of line-item budget controls. For example, under the French system, controls on the line-item composition of program expenditure have been totally abolished with one (important) exception – personnel expenses. It is not permitted to shift money from other input categories to personnel expenses (although ministries can shift money *from* personnel expenses to other input categories.

[21] Very detailed input controls have, moreover, a downside unrelated to performance budgeting issues – namely, that they tend to aggravate considerably the problem of under-execution of budgets which can be seen in many countries. Under-execution means that actual expenditure falls well short of the level of expenditure approved in the budget. Detailed line-item controls contribute to this because, when there is under-spending in one line-item category (e.g. internal training), it is not possible to use the funds approved for other purposes.

The key practical question is then *how much line-item deregulation, and how fast.* It should not be thought that program budgeting demands the complete abolition of central controls on the input allocation of program expenditure. In the French system, for example, control over the allocation to personnel expenses is controlled for the very good reason that France is a traditional civil service system where, once someone is hired, it is very hard to terminate their employment. So if a ministry was to hire more people by shifting funds within its budget, it would be creating commitments to future expenditure which could potentially greatly reduce the budgetary flexibility of the government in the future.

In nearly all program budgeting systems, capital expenditure is protected – that is, spending ministries cannot shift money from capital to current expenditure within programs. The rationale for this is a concern about the temptation ministries would otherwise face to make short-sighted cuts to capital spending to ease their current budget constraints.

Particularly in developing countries, there are other types of line-item controls which it may make sense to retain, such as:

- Line-items for utility bills and other essential payments: this would be relevant if there was a concern that, without such controls, ministries would spend their budgets on other items, leaving insufficient to pay their electricity and phone bills, and would then demand supplementary funding from the MoF to pay these bills.
- Controls over travel, consultancy and other items particularly susceptible to abuse.

These examples highlight the more general point that the practical scope for line-item deregulation in any country depends significantly on the quality of governance and on the degree of civil service discipline. Rather than simply following the extreme examples of France or other countries which have reduced their line-item controls to a minimum, each country needs to make its own case-by-case decisions about what line-item controls to retain. It will also typically make sense to adopt a gradual approach to line-item deregulation.

This being said, it makes no sense at all for any country adopting program budgeting to retain the highly detailed traditional line-item controls referred to above. Most such detailed controls serve no useful function whatsoever. The best approach will be to accompany the move to program budgeting with a review of line-item controls which aims to abolish most detailed controls, and to retain only those for which a clear rationale can be put forward.

Key Readings

Robinson, M. and H. van Eden (2007), "Program Classification", in Robinson (ed.) *Performance Budgeting.*

10. Program Accounting and Costing [22]

This section examines the accounting, information technology (IT) and budget classification requirements of program budgeting. The principal point is that if program budgeting is to work, programs must be integrated into the accounting system, the way in which the budget is structured, and the IT systems which support the budgeting process.

If a government is going to budget by programs, the budget documents and the accounting system must be program friendly. In particular, it must be possible to follow budget execution in programmatic terms – that is, to monitor expenditure program by program during the year so as to make sure that program expenditure authorizations are not exceeded. Insofar as line-item authorizations (for spending by economic classification – see further below) are expressed in program terms, it must be also possible to monitor these. For example, if limits are set on salary expenditure *for each program*, then the accounting system must continuously monitor salary expenditure by program. It is not sufficient, for example, to make an estimate of program expenditure at the end of the year by some approximate means (which would be enough if programs were intended only for reporting and not budget execution purposes – that is, to provide information on the level of expenditure by policy objective).

Budget Classification and Chart of Accounts

For a program budgeting system to work, the "Budget Classification" (BC) and "Chart of Accounts" (COA) must incorporate programs. The BC and COA are classification systems, in which there is a code number for each transaction which summarizes all of its characteristics according to the classification system. These systems cover not only expenditure, but other transactions such as revenue. But our focus here is exclusively on the classification of expenditure. Expenditure on a particular item (say a salary payment, or a purchase of specific supplies) will in these systems be coded with a sequence of numbers such as 1321-325-257-3, where "13" stands for the ministry, "2" for the department within the ministry, "1" for the unit within the department, ""325" the three levels of the economic classification (i.e. the type of input – whether, for example, salary, capital expenditure etc), "257" symbolizes the three levels of the functional classification (more on this later) and the "2" the fund source (e.g. donor funded). The classification categories used are not set in concrete, but may vary between countries. Some countries might have additional ones, such as a geographic classification, in

[22] This section relates to session 2.4 in the generalist course, and session 1.5.4 in the specialist course.

which case there would be an additional code representing the region in which the expenditure took place.

The COA is a coding system used when recording transactions *in the accounting system*. This means that for every payment made (or commitment made), ministry accounting staff would enter the appropriate code when recording the transaction in the accounts (concretely, in the computerized accounting system). The BC is the classification used in the budget itself – that is, in describing in the budget documents the expenditure which is to take place in the coming year. [23] The COA and BC should be integrated in the sense that the former should subsume the latter. In other words, the accounting system should record all the expenditure transaction categories that are the basis of budgetary approvals or information, so as to enable expenditure control and reporting in terms of those categories.

Under program budgeting, programs must be incorporated into the COA and BC. Not only the program, but the subordinate elements of the program hierarchy (sub-programs etc), should be coded for. So if there is a three-level program hierarchy, three digits will be required for coding purposes. It will also be necessary that training and validation processes are put in place to ensure that accounting staff record the program properly when entering expenditure transactions into the accounting system.

Financial Management Information Systems

Financial management information systems (FMIS) refer to computerized systems for managing budgeting and associated processes. Computerized accounting systems are the most basic type of FMIS. However, many other processes are also computerized. Automated expenditure control is an example of such a process, and refers to computerized systems which are used to authorize commitments and/or payments by verifying that they are consistent with the expenditure limits set in the budget. Other computerized system will cover other budget execution functions such as cash management, debt management etc. There are also systems which cover various aspects of budget execution.

Not only must the accounting system be program friendly, but other systems within the FMIS must also be adjusted to work with programs if program budgeting is to succeed. The

[23] Whether for the purpose of describing "control totals" (see the program accounting section) or providing supplementary information on the expected composition of expenditure (e.g. details of the expected breakdown of program expenditure by line item, even when this is indicative rather than binding)

expenditure control system is an obvious example – if the budget sets limits for expenditure by program, it will be necessary that the expenditure control system *as well as the accounting system* operates on a program basis. Concretely, this means that the expenditure control system would prevent commitments being made for spending on program X in excess of the amount voted by parliament for program X. If there are program transfer provisions, the expenditure control system will also need to manage this – it should permit (subject to appropriate authorization) transfers between programs up to the established limit (e.g. five percent), but not in excess of this. If there is a computerized budget preparation system, this also will need to change to recognize the fact that ministries will now be putting forward spending proposals *linked to specific programs*.

The acronym IFMIS refers to an *integrated* financial management information system. This means a single system which to a greater or lesser degree integrates multiple functions. With an IFMIS, accounting, expenditure control, payments, budget preparation and a range of other functions will be built into the same large computerized system. An IFMIS differs from a system which is based on a series of separate systems which draw data from one another but are not otherwise integrated – in the sense that they do not automatically "talk to" one another, so that data entered into one system does not automatically modify relevant variables in other systems. [24]

The scope of an IFMIS – that is, the number of different budgeting and related functions which it incorporates – is not fixed, but varies considerably internationally. To be considered an IFMIS, a system must integrate at least a minimum set of "core" functions (typically including accounting and reports, expenditure control, payments). It may also include certain "non-core" functions, such as budget preparation, cash management, debt management and even performance management (i.e. recording of program objectives, indicators, monitoring against targets).

The key issue for program budgeting is how to ensure that the financial management information systems are program-friendly. Some assume that this requires the introduction of an IFMIS which incorporates all of the functions relevant to program budgeting, including for example budget preparation on a program basis and performance management.

[24] For example, there may be a distinct payments system which will be separate from the accounting system but which will use data from the expenditure control system to determine whether a specific payment is based on a commitment which was previously authorized by the expenditure control system as compatible with budgetary limits.

This may be an inappropriate assumption for a number of reasons. Firstly, if program budgeting is to be introduced in a country which lacks a functioning IFMIS – that is, a country which has a financial management information system based on a set of separate interfaced systems – it needs to be recognized that the replacement of such "legacy systems" (as they are often called) with an IFMIS is not something which can be done overnight. IFMIS implementation takes, as a rough generalization, a minimum of five years. Many countries, when introducing program budgeting systems, wish to move more quickly than this. If they do, this will require them to modify the legacy systems to make them program compatible, even if in the medium of longer term they are planning to introduce an IFMIS. In France, the program budgeting reforms which came into full operation in 2006 were in their early years made to work with the pre-existing financial management information system, modified in the most essential ways. This was despite the fact that even at that time a massive project was already underway for the introduction of the highly-integrated "Chorus" IFMIS system. This is because the Chorus was such a major project that it required many years preparation and implementation, and to wait until this was completed would have inappropriately delayed the introduction of the performance budgeting system.

Secondly, there is a close association between the comprehensiveness of the planned system and the scale of the implementation difficulties and delays. As DFID (2001: 61) note, "international expenditure suggests that there is a real risk of failure in the introduction of large management information systems". It is therefore a big mistake to assume that an IFMIS should necessarily try to integrate everything into one gigantic system. Diamond and Khemani (2005: 3-4) also warn:

> *…the term "integrated financial management information system" can sometimes be erroneously interpreted as describing a system that can capture all the functional processes, and the relevant financial flows, within public expenditure management. However, the complexity of information systems within the government sector is, to a large extent, due to the multiplicity of functions and policy areas. In many functional areas specialized information systems are in place and will still be required even with the implementation of an FMIS.*

French experience once again highlights the problem. It was just as well that performance budgeting was not delayed to await Chorus, because Chorus implementation subsequently

experienced massive problems, to the extent that for a period in 2010 payments to suppliers largely stopped and major payment arrears accumulated.

As noted in the section of performance information systems, it is also wrong to assume that program budgeting requires that a performance management module be integrated into the IFMIS. Program objectives, indicators and targets can be perfectly well managed via a much simpler separate platform which might even be based on Excel spreadsheets. More generally, a move to program budgeting should not be seen as requiring that one over-complicate the computerized financial management system in a way which is only likely to cause major implementation difficulties.

Program Cost Allocation

We now turn to a quite different issue: the methodology guiding the allocation of expenditure between programs. Expenditure is carried out largely by the purchase of inputs [25], so the accounting for expenditure by program requires that accounting staff be able to determine the program under which the expenditure on any given input should be recorded in the accounts. There are two distinctly different aspects of this issue, concerning respectively *direct* and *indirect* costs.

In costing programs, a direct cost is expenditure on an input which contributes to that program *and that program alone*. The same is true in respect to sub-programs – a direct cost is a cost associated with only the sub-program concerned. Assume, for example, that there is an agricultural "extension" service (i.e. a service providing technical advice to farmers) and that this service constitutes a sub-program within, say, a crop industries program. The wages and other employment costs of field advisory staff that spend 100 percent of their time providing advice to farmers would be a direct cost of that sub-program. So would the operating costs of any motor vehicles dedicated exclusively to that service. On the other hand, an indirect cost would in this context be any item of expenditure which contributes partly to that sub-program, but also to other sub-programs. For example, if a motor vehicle were sometimes used to support agricultural extension service, and sometimes to support staff delivering other types of services within the crop industries program (or even other programs), it would be an indirect cost.

[25] Of course, there are also "transfer payments", such as income support benefits.

Consider this issue from the perspective of accounting for expenditure between programs within a specific ministry. There are within any ministry a range of "overheads" – expenditures on services which support the ministry as a whole. Such overheads include, for example, the employment costs of top ministry management, the function of which is to manage across the ministry as a whole. They also include the vital service functions within the department such as finance and human resources. If we assume for the moment that there is no separate "administration" program covering these overheads, then they are by definition indirect costs which would have to be allocated to the ministry's programs. For example, in the case of the education ministry, it would be necessary to determine how much of the ministry's finance, human resources, top management etc expenditure contributed to the primary school education program, how much to the secondary education program, and so on.

Direct costs are, in principle, the "easy" part of program costing. Because they are costs which contribute to only one program (or sub-program), all that has to be done is to ensure that they are recorded against the correct program or sub-program. (As noted below, however, even this can be a significant challenge in some developing countries.)

The allocation of indirect costs is less straightforward. Take the example of the regional office staff of an environmental agency, many of whom might be involved in a wide variety of different services, ranging perhaps from nature protection to environmental pollution management. What methods might be used to allocate their employment costs to specific outputs?

One method would be to require the staff concerned to fill in time sheets recording how their time is allocated, in the manner that law firms typically use for billing purposes. Government departments in some countries have implemented such time recording systems. This may be, however, a costly process, particularly if the time sheets are to be reasonably accurate (which will require careful monitoring by management to make sure that staff fill in their time sheets accurately). This points to a key issue which arises in allocating some indirect costs: the expense of the cost allocation process itself. Because of the expense of acquiring truly accurate cost allocations, it is commonplace for shortcuts to be adopted, as a conscious compromise between accuracy and expense. In the case of staff time allocations, for example, the compromise might be to require staff to fill in time sheets for a specific period during the year (say one month) and then assume that this sample data was representative of the whole year.

Information technology has in some cases greatly reduced the costs of obtaining accurate allocations of indirect costs. Take the examples of telephone costs or electricity usage. It is a relatively straightforward and cheap matter these days to generate accurate information apportioning an organization's total phone bill to individual phone lines, or electricity usage amongst different organizational units. In the past, however, this was impossible or very difficult. Note, however, that in some developing countries, technology to record the use of such services may not be affordable, and as a consequence such utilities remain in effect indirect services.

Indirect cost allocation is a task of *management accounting* which, while widely employed in the private sector, is not typically part of the arsenal of accounting tools of government. It involves the definition and application of "cost drivers" to allocate expenditure. A cost driver is a formula or principle which is used to determine the proportions in which an indirect cost is split between two or more programs (or sub-programs) to which it contributes. In the example above, both the apportionment of telephone costs of a pro-rated basis reflecting direct service staff, and the apportionment is accordance with computerized usage records, represent (different) cost drivers. One may be arbitrary and the other accurate, but they are both cost drivers. Because cost drivers provide the basis for determining the proportions in which indirect costs are allocated, the term *allocation basis* is sometimes used to mean the same thing as a cost driver.

To keep things simple, accountants often allocate indirect costs using very simple cost drivers. For example, the cost allocation formula might assume that all the support services in the ministry contribute to programs in a manner proportionate to the staff numbers within the major organizational units in the ministry. For example, if thirty percent of the staff of the ministry of education work directly on the primary education program, it might be assumed that thirty percent of the cost of the ministry's support services is attributable to the primary education program. However, this would be a very crude approximation which might be very inaccurate.

To allocate indirect costs in a reasonably accurate (as opposed to highly arbitrary) way is a complex business, requiring not only additional highly skilled accountants, but more sophisticated IT systems. Amongst the sophisticated indirect cost allocation methods available is *activity based costing* (ABC). There are two defining features of ABC. The first is that it is a two-stage costing process as applied to program costing, in which inputs are allocated firstly to

"activities" (including support activities such as HR services or IT services) and activity costs are then allocated between programs. The second defining feature of ABC is an insistence upon the use of cost drivers which are not arbitrary but which provide an accurate reflection of the contribution which inputs and activities actually make to outputs. Advocates of AOB view it as requiring the use of cost drivers which constitute 'a model of organizational resource consumption'. The ABC emphasis upon accurate cost drivers is laudable. However, the price of accuracy can in many instances be unacceptably high. Even in developed countries it is not clear that ABC has been a success in the public sector. In developing countries, it is generally speaking far too demanding of skilled human resources and sophisticated IT systems to be feasible or justifiable in cost terms.

Administration Programs

The difficulty of accurate indirect cost allocation has led, for most government, to a pragmatic modification of the program budgeting principle. This is the use of distinct "administration" programs to cover the overhead (support service) of each ministry. Administration programs are, clearly, inconsistent with the pure program budgeting principle of outcome/output based programs, as outlined in Section 8. Internal support services do not provide services to external clients, and are therefore by definition not outputs (see Section 3). Moreover, precisely because they do not deliver services to the public, support services are not capable in isolation of achieving outcomes for the community. Rather, they support those parts of the ministry which do provide services directly to the public, and which through those services achieve outcomes. Therefore, if all programs were defined in terms of outputs and outcomes, there would be no administration programs and all ministry-wide overheads would be indirect costs which will need to be allocated between programs.

However, precisely because the allocation of indirect costs with an acceptable level of accuracy is a demanding and technically complex undertaking which is not cheap, the majority of countries choose to largely avoid it through the use of administration programs (also known as "management", "support" or "corporate services" programs) [26]. Administration programs cover the costs of ministry support services which would otherwise be difficult to allocate. Countries which make use of administration programs take the view, however, that it is better to compromise on the program budgeting principle in this way than to use of an arbitrary cost

[26] And, analogously, they commonly make use of administration sub-programs within each program, covering program-wide overheads.

allocation approach which would yield “full cost” program expenditure measures which might be highly inaccurate.

If the only justification for administration programs was the difficulty of accurate indirect cost allocation, it might be appropriate to view administration programs as a temporary measure. In other words, administration programs could be viewed as no more than a practical mechanism to be used in the early days of a program budgeting system in countries where the capacity and systems for accurate allocation of indirect costs do not yet exists. In the longer term, the aspiration would be phase out management programs as and when the capacity to allocate indirect costs is developed. This is a view which many people take.

It is, however, possible to argue the use of management programs to cover ministry support services can be justified as a matter of principle even in countries which have the technical capacity to allocate expenditure been programs.

To see why, keep in mind that program budgets are in most countries expenditure limits (“control totals”), and not simply accounting measures of the cost of producing specific “product lines”. In other words, as discussed in Section 9, when the parliament approves a program budget, it is telling ministries that they are to spend no more than $x on program A, $y on program B, etc. This means that if the expenditure of the ministry’s support services is included in its outcome/output-based programs, the government is in effect instructing the ministry about the extent of the support the ministry’s support services should provide to each of the ministry’s output-based programs.

Suppose, for example, the education ministry has only two programs – primary and secondary education – and no management program. Imagine, moreover, that the ministry’s organizational structure consists of a primary school department and a secondary school department – each of which is budgeted internally by the ministry to receive 40% of the ministry’s overall budget – and a support services department which is budgeted internally to receive 20% of the ministry’s budget. Supposed further that it is expected at the time the budget is prepared that the support services department will provide equal levels of support to each of the two main departments. Reflecting this, the program budget approved by parliament will allocated one-half of the ministry’s budget to each of the primary and secondary education programs, in each case including that program’s share of expenditure on support services.

To ensure that neither program breeches it expenditure authorization, the education ministry will then need to make sure that the support services department does not devote substantially more than one-half of its expenditure to supporting either of the two ministry programs. More generally, because support services expenditure is included in the output based programs, it becomes necessary in budget execution for the support services department to rigorously manage the balance of support it provided to the other ministry departments.

Herein lies the problem. Why would the political leadership or the MoF wish to dictate the way in which education ministry support services allocate their efforts between supporting primary and secondary education during the year? If, for example, unexpected events mean that the primary school program has a greater-than-anticipated requirement for the support of the HR group in staff recruitment during the year, shouldn't the ministry be left with the flexibility to shift the disposition of its support services accordingly?

If one believes that ministries should retain the flexibility to allocate support services to where they are needed during the year, then it makes sense *in principle* to have a management program. That management program should included all those ministry support services which are generic and which can be flexibly reallocated during the year from supporting one output-based program to another to meet shifting support requirements [27].

Of course, it is still useful to know ex post where the support services are allocated – in order, ideally, to have the best possible accounting measures of the full costs of delivering each "product line". For this purpose, additional cost analysis can be carried out to allocate the management program expenditure between the output-based programs.

This is, incidentally, exactly what the French do – even though France is perfectly capable of carrying out the accounting allocation of support service expenditure between output-based programs, it nevertheless makes us of management ("support") programs and supplements this with additional cost analysis.

[27] It is relevant in this context that in Australia, where there are no administration programs, programs are not used by parliament as control totals (legal appropriation categories). This means that the allocation of funds to the budget is purely indicative, and therefore does not in any way constrain the actual allocation of support services between the main program areas of ministries.

Salary Costs and Programs

It was suggested above that the allocation of *direct* cost is a relatively easy task. In many developing countries, even direct cost allocation between programs can present significant challenges. It is not uncommon, for example, for ministry accounting systems in such countries not to differentiate personnel expenditures by ministry organizational unit. This problem is often compounded by systems in which, when staff from one ministry are posted to another ministry, they continue to be paid by their original ministry, as a result of which their salaries are not recorded as expenditure of the ministry in which they are actually working. This means that the accounting system in, say, the education ministry might give no indication of the total wage costs of staff working on primary education, making it essentially impossible to cost a primary education program.

For these reasons, developing countries are often tempted to record all ministry personnel expenditure under the "administration" program. While one can sympathize with the practical reasons for doing this, such a course of action fundamentally undermines the logic of program budgeting. Given that personnel costs represent a large portion of government expenditure in most countries, the results of excluding such expenditure (at least on the current expenditure side) is "programs" which only record minor expenditure (e.g. furniture, supplies) and which are therefore of very little use as a tool of expenditure prioritization.

In countries which face this problem, the proper allocation of salary costs to relevant programs should be the first priority on the accounting side in preparing for the introduction of a program budgeting system. All personnel charges of staff members working exclusively on specific programs (which will account for most personnel expenses) should be allocated to programs, with only the salaries of staff in support services included in the administration program.

Cost Allocation and the Program Hierarchy

The cost attribution issue – in respect not only of indirect but also of direct costs – is relevant to the decision of how many levels to have in the program hierarchy. The accounting system must be capable of recording expenditure against each element in the program hierarchy. This means that for each additional level added to the program hierarchy, the cost attribution issue becomes potentially more challenging. If, for example, you have a program structure comprised of four levels, it is necessary to be able to follow budget execution – that is, to monitor expenditure via the accounting system during the year – right down to the lowest of those four levels. Suppose that each program is comprised of four sub-programs, each sub-program of four

sub-sub-programs and, finally, each sub-sub-program or four sub-sub-sub-programs. If this is the case, then a four level program structure involves classifying and monitoring expenditure in sixteen times as many categories as does a two-level structure. This might greatly increase the challenge of correctly recording expenditure as well as that of allocating indirect costs. Given the magnitude of the difficulties many developing countries face with their accounting systems, it will hardly come as a surprise to say that few succeed in successfully implementing program hierarchies with three or four levels.

The simplest option – one which may be appropriate for many developing countries, for example – is to have a program hierarchy with only one or two levels (programs, or programs and sub-programs). The appropriate approach to this issue will also need to take into account the relationship between organizational units and program elements, which is discussed below.

Program Cost Estimation in Budget Preparation

The third and final issue addressed in this section concerns the estimation of program costs in budget preparation. In other words, during the preparation of the budget, it will be necessary to put a figure on each program's expenditure for the coming year. [28] Each ministry will therefore need some method to make the estimate of planned expenditure on each program which it includes in its "budget bid" presented to the MoF.

The simplest and most common approach to estimating program costs in budget preparation is to take the previous year's expenditure on the program and adjust it for (1) input cost changes or demand changes; and (2) policy changes such as an expansion of the program or a change in program design.

It is sometimes suggested, by contrast, that under a program budgeting system – or under a version of it which some call "output budgeting" – it is essential that the expenditure estimate for each program be calculated by taking the planned outputs to be produced by the program and multiplying them by the unit cost of the outputs concerned. [29] In other words, unit costs are supposed to be used in budget preparation to calculate program expenditure requirements as a function of the quantity of services to be delivered to the public. Under this approach, the program classification of expenditure would extend beyond programs and sub-programs to the

[28] As well, ideally, as the following three or so years, in a medium-term budgeting framework – see Section 18.
[29] Or, as a variation on that idea, as a calculation based on the number of planned program activities times the cost of those activities (what is sometimes called "activity-based budgeting").

level of the specific individual types of outputs which comprise each sub-program, and budget estimation would proceed in a bottom-up process based on unit costs and planned quantities of each output.

This is essentially the application to program costing of the idea of formula funding, whether based on outputs or activities. As discussed in subsequent sections on formula funding and purchaser-provider systems, unit costs are a powerful tool when selectively applied to the right types of public services. The most outstanding international example of a sectoral performance budgeting system based on unit costs in the "diagnostic related group" (DRG) hospital funding system. In education, where costs per student at particular levels of schooling tend to be relatively standard, unit costs are also a useful budgeting and performance management tool.

However, unit costs can never be an across-the-board budgeting instrument, because there are many public services which do not have a stable unit cost. This is because of "heterogeneity"—average costs vary because of differences in the effort required because of the circumstances of particular cases and also because of what we in Section 12 as "contingent capacity services". Because many government outputs do not have stable unit costs, it is quite inappropriate to seek to use unit costs as the method for estimating the required budgets or all programs across government in a program budgeting system.

Moreover, the "output budgeting" version of program budgeting fails to recognize the enormous complexity of calculating the unit costs each of the large number of types of services delivered by government.

At its worst, the application of the unit cost approach to program costing can degenerate into a justification of program expenditure in terms of planned quantities and unit costs of internal *activities* such as meetings, internal training seminars, and the production of ministry reports. Such an approach confuses activities with outputs, and activity-based budgeting with output-based budgeting. A focus on activities (work processes) is far removed from the performance budgeting emphasis on "outputs" (services delivered to the public) and outcomes (benefits realized by the community).

Key Readings

Robinson, M. (2007), "Cost Information" in M. Robinson (ed.) *Performance Budgeting: Linking Funding and Results*.

Diamond, J. and P. Khemani (2005), *Introducing Financial Management Information Systems in Developing Countries*, IMF Working Paper, obtainable at http://blog-pfm.imf.org/pfmblog/2008/01/introducing-fin.html.

Khan, A. and M. Pessoa (2009), "Conceptual Design: A Critical Element of a Successful Government Financial Management Information System Project", IMF, PFM Technical Manual, obtainable at http://blog-pfm.imf.org/pfmblog/fad-technical-notes-and-manuals-on-public-financial-management.html.

Peterson, S. (2006), *Automating Public Financial Management in Developing Countries*, John F. Kennedy School of Government, Harvard University, obtainable at http://web.hks.harvard.edu/publications/workingpapers/Faculty_Name.aspx?PersonId=213

Robinson, M. and D. Last (2009), *A Basic Model of Performance-Based Budgeting*, IMF Technical Manual 1, obtainable at http://blog-pfm.imf.org/pfmblog/fad-technical-notes-and-manuals-on-public-financial-management.html.

Wynne, A. (2005), *Public Financial Management Reforms in Developing Countries: Lessons of Experience from Ghana, Tanzania and Uganda*, African Capacity Building Foundation Working Paper no. 7, obtainable at http://www.acbf-pact.org/knowledge/documents/Working%20Paper%207%20on%20Financial%20Management-Final.pdf.

11. Special Topics in Program Budgeting

This section starts with a discussion of the relationship between program structure and organizational structure, followed by an examination of the closely-related question of the distinctive design features of program structures in developing countries. The relationship of programs to the functional classification of the budget is then considered. The final topic is the relationship of program structures of different levels of government (in particular, national and subnational government).

Programs and Organizational Structure

As discussed in previous sections, the program budgeting ideal is that programs should be results-based. That is, programs should be based on groups of services delivered to external parties ("outputs") which have common outcomes – "product lines" in the shorthand terminology of the section on program budgeting fundamentals. It is by basing programs on results in this manner that programs can serve as a useful tool for expenditure prioritization and increasing pressure upon ministries to perform.

However, organizational structures are not always results-based. In most government structures, some organizational units exist which are not defined on the basis of a "product line". Differences between program structure and organizational structures can be a problem because program budgeting will be most effective when there is clear organizational responsibility for program management – that is, when a single organizational unit can be held responsible for the results delivered by a program, and exercises the budgetary flexibilities which program budgeting provides.

The question therefore arises as to whether, to what degree, and in what manner, program structure and organizational structure can be brought together.

An organizational structure based on "product lines" would mean that all or most organizational units would be defined in terms of distinctive categories of service delivered to the public, with all services of that type delivered by the same organizational unit. Actual government organizational structures differ from this in a variety of ways, including:

- Organizational structures which are defined in terms of a specific client group (specific industries, regions, types of individuals) and not in terms of a specific category of outputs, and which might therefore provide a quite divergent range of outputs to their target client group rather than the common product line. For example, agriculture ministries in certain countries have taken on responsibility for HIV/AIDS prevention activities targeted at farmers, on the theory that the agriculture ministry's network of contacts with the farming community places them in a good position to manage this service.
- Organizational structures based on functional lines – for example, based on specific professional competences and types of activity (e.g. an organizational structure for the ministry of public works based on separate engineering, planning etc department, as opposed to a structure based on the different types of infrastructure which the ministry builds).
- As a special case of the former category, support service units within ministries or agencies extending their role to the provision of services to the public, and in doing so duplicate the role of other ministries or agencies. A classic example of this is the army medical corps – the original role of which is to provide medical services to the army itself – becoming a significant provider of health treatment services to the public, in "competition" with the health ministry.
- Support service organizational units: organizational units within ministries which are defined in terms of the provision of specific types of support services (e.g. human resources management, IT, finance). These were discussed in the previous section of this manual.

Organizational units which contribute to several quite different product lines are particularly problematic from a program budgeting point of view. Less difficult to resolve is the alternative situation where multiple organizational units contribute to a specific product line, without there being a person in hierarchical authority over all of those organizational units who is able to provide the integrated management which a program comprising all of those units would require.

In a broader "managing for results" (MFR) approach to public sector reform, it often makes sense to undertake restructuring to place organizational structures on more of a "product line" basis than one tends to see in traditional civil services. For example, it could be appropriate to phase out the role of the army in health services to the public, with the health ministry taking full responsibility for this. And if there are two or more organizational units with a single ministry which are involved in delivering the same "product line", the units may either be

merged or someone can be appointed who has hierarchical authority over all of the units concerned. Functional structures should be, in many cases, replaced with product line structures. Such reorganization can be expected to considerably reinforce the results focus which program budgeting aims to encourage.

Possible organizational restructuring alone cannot be expected to totally resolve the issue of conflicts between organizational and program structure, for two reasons.

The first reason is that, even if a more results-oriented organizational structure is highly desirable, it may not be feasible or even desirable to extensively modify organizational structures at the same time as moving to program budgeting. The magnitude of such a restructuring task may be too great, or it may simply be impossible to obtain adequate support for organizational restructuring at that time. An explicit decision may under such circumstances be made to leave organizational restructuring to a later stage in the overall public sector reform effort.

The second reason is that a *purely* "product line" organizational structure is not desirable. Under a thorough-going MFR reform, it may well be that some ministry-wide support and coordination service functions are taken over by the service delivery units of the ministry. For example, MFR argues for greater managerial flexibility over human resources (hiring, remuneration etc), and insofar as this occurs the power of ministry human resources directorates (and civil service ministries) will be significantly reduced, with some of their functions transferred to front-line organizational units. Nevertheless, in any rational organizational structure it will always make sense for a core of support services to be provided centrally by functional units, rather than completely split between the service delivery units of the ministry. There will always be a ministry finance unit, and usually some type of IT unit, HR unit etc. Some ministries will also need to maintain regional offices which, for size reasons, deliver multiple products rather than specializing in single product lines. So even if one wishes to apply the "product line" organizational principle to the maximum degree to align programs and structures, one cannot expect to achieve full alignment by this means.

This means that when introducing program budgeting, some conflicts between organizational structure and a "pure" program structure based entirely on results are inevitable. Some of these conflicts may be temporary, in the sense that they are capable of future resolution via organizational restructuring. Some, however, will be permanent.

There are two approaches which can be taken to this conflict between organizational structure and programs. The first is to define programs around organizational structures, even when this involves significant departures from the principle of results-based programs. The second is to maintain the principle of results-based programs and to accept that there will therefore be difference between organizational structure and program structure – some of which may be temporary, but others of which will be permanent. Management strategies will be put into place to address the conflict between program and organization structure.

In practice, some mix of these two approaches is not only usual, but to some degree inevitable.

Ministry Boundaries and Internal Support Services

There are two broad areas where there is wide agreement that it is appropriate that there be some compromise of the program budgeting principle by adapting programs to organizational structure.

The first of these is in respect to ministry organizational boundaries. The great majority of countries observe the principle that *programs should not cross ministry boundaries*. It might appear attractive to say that if two different ministries are contributing to the same "product line", then a cross-ministry program should be established based on that product line. For example, the idea of a single "AIDS" program grouping together all of the HIV/AIDS prevention and treatment activities of multiple ministries sounds very appealing. Against this, however, is the fact that it is essential that each ministry has its own clear budget. It would be completely unworkable for the budget to simply allocate funds to programs shared between two or more ministries and then leave those ministries to determine how those funds were to be split between them. If there were to be multi-ministry programs, it would therefore be essential for the budget to clearly indicate how much of the program's allocation would be directed to each of the ministries involved. Essentially, this would involve having sub-programs which were limited to single ministries, which would essentially take us back to the principle that programs (or in this case sub-program) should not cross ministry boundaries. [30]

[30] The allocation of multi-ministry programs between ministries at the sub-program level can in principle work. But it would have the undesirable consequence of forcing the parliament to approve the allocation of the budget at the relatively detailed sub-program level rather than, as is the usual international practice, at the broad program level.

Some countries attempt to bring greater coordination to areas of related service delivery by designating super groupings of programs. The best example of this is France, where programs are grouped together into large *missions*, some of which cross ministerial boundaries. The principle that programs remain within ministry boundaries nevertheless continues to be respected.

Cross-Ministry Programs in Africa

In certain African countries which have introduced program classifications in their budgets, there has been a deliberate decision to introduce programs which are shared by two or more ministries belonging to the same sector. To make this work, the legal appropriations in the budget specifying the shares each of the ministries concerned has of the overall budget of the program. The context of this approach is government structures characterized by a far larger number of ministries than exist in most countries around the world – for example, as many as 50 ministries. Under these circumstances, it is often the case that multiple ministries are involved in pursuing the same outcomes. To ensure co-ordination, these countries not only employ sector-based programs, but place much emphasis on sector-wide policy coordination. This is understandable, but a first-best solution to the programs of fragmentation of the policy effort between too many small ministries would be to radically reduce the number of ministries. In many cases, the reason for the multiplication of ministries is the desire to create a larger number of ministerial positions for politicians.

If programs are to be kept within ministry boundaries, one thing which immediately follows is there will need to be programs which are exclusive to coordinating ministries, and that these programs will for the most part not be defined in terms of outputs and outcomes delivered to the public, but in terms of the type of support service provided to the government as a whole. For example, a civil service ministry or commission might have a "Civil Service" program with an overarching program objective defined along the lines of "a quality civil service able to provide excellent policy advice and service delivery". Clearly this objective is not an outcome, but it is an important means towards delivering outcomes to the community.

The second area where the majority of countries with program budgeting systems compromise the pure program budgeting principle is in establishing an administration program within each ministry, which – as discussed in Section 10 and elsewhere above – groups together all of the support services and other "overhead" costs of the ministry. They do this in order to avoid the complexities of the alternative approach, which requires that the costs of all support services be attributed to the results-based programs which they support. As discussed in the section on program accounting and costing, this alternative approach necessary involves quite complex

indirect cost allocation, which for many countries is either beyond their technical capacity or which they cannot reasonably afford to undertake. Only a minority of OECD countries – Australia being one example – have taken this latter approach.

Administration programs will in general have program objectives which do not refer to outcomes. For example, suppose there is a government agency which has a program which manages the provision of office accommodation to government ministries. The objective of such a program would be something like "ensuring that the needs of government ministries agencies for appropriate premises are met in a timely fashion". In other words, the objective would refer only to the output being provided. This is because the service involved is not directly responsible for delivering outcomes.

Broader Conflicts between Program and Organizational Structure

What about conflicts between organizational and program structure which go beyond these obvious areas for compromise just discussed? Here, the dangers of further compromise of the program budgeting principle are much greater. If one simply modifies the program structure wherever it conflicts with organizational structure, one ends up not with a program budget but an organizational unit budget. The traditional budgeting model in quite a few countries (including most of the "Latin" countries, South America, non-Anglophone Africa, and much of eastern Europe) was one in which budget appropriations were made not only by line-item (economic classification) but also to internal ministry organizational units. If the program principle is to be simply made to fit the organizational structure, the result will be something just like this type of traditional structure. And insofar as organizational structure is not result-based, the fundamental objective of making budgeting more results-focused will be undermined.

If the program budgeting principle is not to be further undermined, the question arises of how to manage conflicts between organizational and program structure. A common approach to this problem – which France exemplifies – is to appoint program managers. When the organizational unit corresponds to a program, the organizational unit and program managers would be the same person. But where organization and programs diverge, they would be different people. The program manager's role would then be to maintain the product-line focus and offset the threat to that focus arising from a functional, sector or other organizational principle. Specifically, the program manager would work with organizational unit managers to maintain

the focus on overarching program objectives and to promote coordination between the various organizational units involved in delivering the program. If this approach works well, one would not need to worry about divergences between program and organizational structure.

This approach has considerable attractions. However, it has to be recognized that it also raises potentially important difficulties – in particular, that the existence of two cross-cutting layers of management can lead all too easily either to continual conflict between managers (which undermines performance) or to the marginalization of one of the two managers (usually the program manager). Conflict between program and organizational unit managers has apparently been experienced in France.

France: Role of the *Responsable de Programme (program manager)* [31]

"The *responsable de programme* is designated by the relevant minister to assure the direction of the program. He collaborates in the formulation of the strategic objectives of the program, under the authority of the minister. He is responsible for the putting into effect of these objectives and their realization. To this end, he translates the strategic objective of the program into operational objectives relevant to each of the organizational units which has a role in the program, by means of dialogue with the managers of those organizational units."

Program Structures in Developing Countries

In the light of the above discussion, it makes sense for developing countries to avoid unnecessary complications in their program structures by:

- Avoiding multi-ministry programs. If it is desired to recognize in the program structure the need for cross-ministry collaboration in overlapping areas of responsibility, it will be better to use the French approach of multi-program missions than to have shared programs.
- Opting for administration programs rather than attempting the onerous task of indirect cost allocation.
- Keeping their program hierarchies simple – in general, limited to just one or two levels (programs, or programs and sub-programs).

In some developing countries, organizational restructuring may be even more important in the medium term. This is particularly the case in countries where, for example, there are far too

[31] Source: France (2004), p. 13.

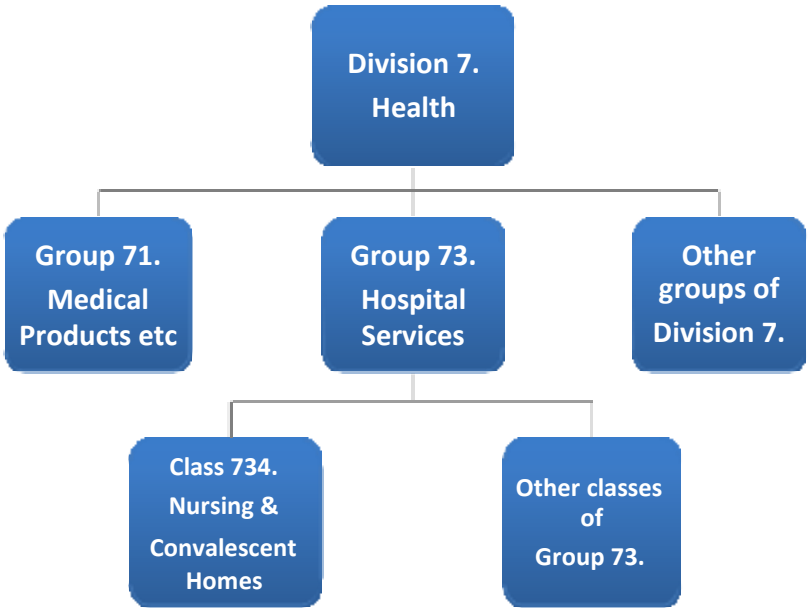
many ministries (sometimes forty or more). Where this is the case, adhering to the principle that programs should not cross ministry boundaries does more damage to the program budgeting principle than in countries with a smaller number of larger ministries. In some developing countries, duplication of functions between ministries can also be a particularly serious problem.

Functions and Programs

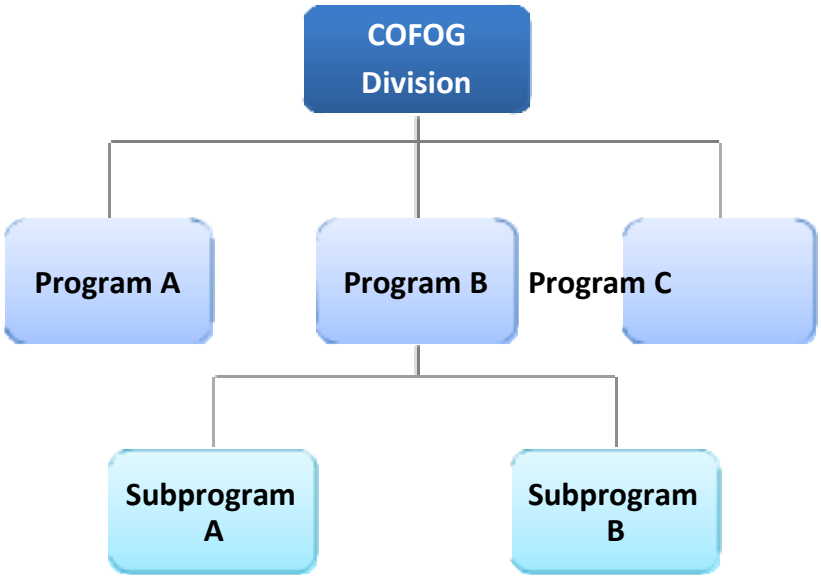
A question which tends to cause considerable confusion is the relationship between program structure and the "functional" classification of government expenditure. The standard international functional classification of the budget is that of the United Nation's COFOG (Classification of the Functions of Government) system. The COFOG classification was designed in order to permit, via a standardized classification of government expenditure, international comparisons of the allocation of resources between policy areas. Note, however, that a number of countries have for historical reasons "functional" classifications which differ somewhat from COFOG.

The COFOG Functional Classification

The COFOG functional classification consists is a hierarchical structure of three levels. The top level ("divisions") consists of four broad categories such as "defense", "public order and safety", "health" and "education". There are ten such divisions. Below this are two lower levels: "groups" and "classes". Thus in the education division, the *groups* include: "pre-primary and primary education", "secondary education", "tertiary education" and "subsidiary services to education". And within, say, the "pre-primary and primary education" group, there two *classes*: "pre-primary" and "primary". Graphically, this hierarchical structure can be represented as follows, taken the example of health (Division 7). (The example shows only some of the groups under the health division, and just one example of the decomposition of groups into classes).



The key reason why there is confusion about the relationship between functions and programs is that the term “function” is somewhat misleading and the majority of the functional categories are really categories of outputs – “product lines”, so to speak. The functional classification is not a pure results-based classification. Nevertheless, the functional classification of expenditure looks so much like a program classification – particularly one with administration programs – that it is natural to wonder why both are needed and whether the two structures should be fused in some way. The most common way of linking the program and functional structures is to insist that programs should all fit within one or other of the ten broad COFOG divisions – with no programs permitted to cross the divisional boundaries. This is illustrated in the graphic below.

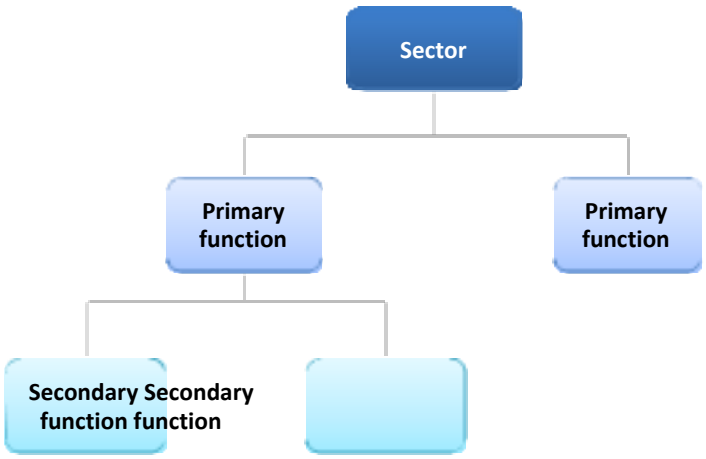


The alternative view is that the functional classification should not constrain the program structure and should not be formally incorporated into it. This point of view is based on the recognition that the functional classification is only intended to be a statistical tool – for international comparisons of resource allocation – and not a budget classification in terms of which the budget should allocate resources.

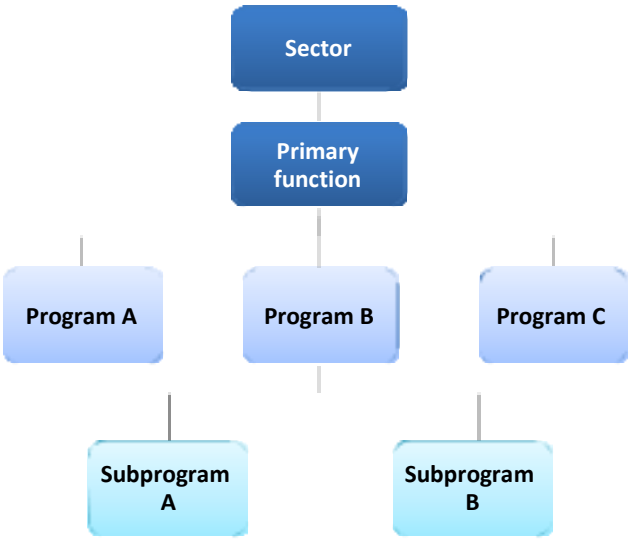
The potential problem with making the program structure fit the functional structure is that the choice of programs in any country should to reflect the specific policy challenges of that country. Thus in a country where the fight against desertification is of crucial importance, it may make sense to have a desertification program, whereas in many other countries no such program will be required. To restrict the program structure to make it fit with a given functional classification reducing flexibility in the definition of programs. This problem is not really a significant one if all that is proposed is to keep programs without the boundaries of the COFOG *divisions* (the top level of the COFOG hierarchy), although even there it is not clear why it is appropriate to ban programs which cross divisional boundaries. However, it is a much greater problem if the program structure is to be made to conform to a more detailed functional categorization, as illustrated by the following case study.

Programs and Functions: a Case Study

One African country which decided to integrate its functional and program classification encountered major problems as a result of the manner in which it chose to do so. The country concerned had in place a pre-COFOG functional classification, which like COFOG was a three level hierarchy, as follows:



In this country's structure, "sectors" (of which there are 10) were similar to COFOG divisions, "primary functions" to COFOG groups; and "secondary functions" to COFOG classes. The fact that the functional classification was somewhat different from COFOG was not, however, the problem. The problem was that it was decided to make programs subordinated to the second level of the functional classification, rather than the top level. This produced the following integrated functional and program structure:



The effect of this was to require that no program should cross the boundaries the "primary" (second level) functions.

The problem with this was that that the primary functions, of which there were approximately seventy, were far too detailed for this purpose – and much more detailed than, say, the ten divisions of the top level of the COFOG classification. Bear in mind that in most countries, the total number of programs across government is of the order of 150-200. Insisting that no program should cross the boundaries of the primary functions proved to be extremely constraining, and made it impossible to create certain programs which appeared important to the government. Following advice, it changed its position, and chose to define the top "function" level of the program structure in terms of the ten COFOG divisions.

Programs and Levels of Government

One question which arises in defining program structures is the relationship between the program structures of different levels of government in a country. Should the national and subnational governments adopt the same program structure? Or should the different levels of government go their own ways in developing their own structures without any regard to what other levels of government are doing?

This is an area where different countries take different approaches. In many federal countries, where subnational government is constitutionally autonomous and is protective of this autonomy, each level of government develops its own structure without paying attention to what other levels of government are doing (this is true, for example, in Australia and Canada).

In many other countries, however, there is a belief that program structures should be in some way standardized or harmonized. This is particularly true in developing countries where there are single national plans which are intended to guide the entire public sector.

Even if one aims for the maximum degree of harmonization, one cannot have exactly the same programs at national and subnational levels. Insofar as the levels of government have different functions, they require different programs. Even in the same service area, their functions may differ, and it might be thought best to reflect this in program titles. For example, it may be that in school education, the national government is responsible for the development of national standards (e.g. a common core curriculum), while subnational government actually delivers the

service. This may lead to the national program being titled, say, “education policy”, while the subnational programs are simply labeled “education”.

Key Readings

Robinson, M and H. Van Eden (2007), “Program Classification”, in M. Robinson (ed.) *Performance Budgeting: Linking Funding and Results*.

Diamond, J (2003) *From Program to Performance Budgeting: The Challenge for Emerging Market Economies*, IMF Working Paper, obtainable at http://www.imf.org/external/pubs/ft/wp/2003/wp03169.pdf.

IMF (2001), *Government Financial Statistics Manual 2001*, Washington: IMF (for detailed information on the COFOG functional classification), obtainable at http://www.imf.org/external/pubs/ft/gfs/manual/

12. Formula Funding and Purchaser Provider Systems

Formula funding – of which the purchaser-provider mechanism is one type -- represents an attempt to create a much tighter linkage between funding and results than characterizes program budgeting. Under program budgeting, performance information is taken into account in determining funding, but there is no automatic linkage. Under formula funding, the linkage between funding and results is a very direct one.

When used as a form of performance budgeting, formula funding means that the funding provided to a government ministry or agency is determined in whole or part as on the basis of an explicit algebraic formula based on outputs or outcomes ("results" for short). The formula might link funding to *expected* (i.e. future) results – as, for example, when school funding is based on the number of students who enroll in a school at the beginning of the year. Or it could be based on the actual results which the agency has delivered – which would be the case if school funding was based on the number of students who actually attend the school during the year (i.e. if a student enrolls at the beginning of the year, but then drops out, no funding would be received). The latter type of formula funding, in which funding depends on results actually delivered, is what is referred to as a purchaser-provider system.

Many formula funding systems are cost-based. For example, if schools are being largely funded by formula based on the number of their students, then it is obviously necessary that the amount of funding per student is based on an estimate of what it should cost to teach one student. However, formulas are also used to provide performance bonus funding which are not based (and could not conceivably be based) on cost estimates. For example, in some public university systems, core funding based on per-student amounts is accompanied by supplementary bonuses linked to outcome measures such as percentages of graduates who find employment. Nobody knows what the cost is of increasing the graduate employment rate by, say, 5 percent, so the level of the bonus which is paid is in a sense arbitrary. This does not, however, undermine its function are providing an incentive to improve quality and achieve better outcomes.

The purchaser-provider mechanism has been used with considerable success in certain sectors of government, most notably hospital services where is took the form of so-called "diagnostic-related group" based funding. [32] Broadly, the idea is that hospitals are paid by the government for the outputs they deliver, and *only* for the outputs they deliver. Each different treatment (e.g. treatment of a hip fracture patient) has a standard price. If it costs the hospital more to provide that treatment than it is paid, it makes a loss. If, on the other hand, it succeeds in

[32] For a detailed description of this system, see Robinson and Brumby (2005), section V.

delivering the service for less than the price paid, the hospital makes a profit. In this way, a powerful incentive in introduced for the enhancement of hospital efficiency (cost-containment). When this system was originally introduced (in the USA), there were enormous concerns that it would lead to serious adverse consequences such as the erosion of the quality of services and the refusal to treat higher-cost patients. In practice, these fears turned out to be greatly exaggerated, in part because of other complementary quality-control systems and in part because of the considerable professional pressures for the maintenance of patient service. Overall, experience has demonstrated the DRG funding system to be a great success in keeping costs down without unacceptable adverse consequences. For this reason, an increasing number of countries have adopted it as the funding model for their public hospitals.

The purchaser-provider system is not, however, something which is suitable for all or most public services. This was demonstrated clearly when Australia and New Zealand attempted during the 1990s and early 2000s to reform their entire budgeting systems to place them on a purchaser-provider basis, under what was known as "accrual output budgeting". The experiment failed badly. [33]

Limits to Formula Funding and Purchaser Provider

Formula funding and purchaser-provider arrangements require a stable relationship between output quantity and cost – expressed differently, a stable output unit cost. These types of funding arrangements therefore suit what have sometime been called "production"-type agencies: that is, agencies which produce large volumes of a limited number of fairly standard services. Motor license testing is a good example. The output delivered by the agency which issues motor licenses to drivers is standard in the sense that every applicant for a license has to undertake the same test. The cost per license issued is therefore stable, and there would be no difficulty in principle in basing funding on numbers of licenses issued.

There are, however, many services produced by government which are not at all like this, and which are unsuited to funding on this type of basis. *Heterogeneous* and *contingent capacity* services require particular mention here.

As discussed earlier in this manual, output heterogeneity arises when the activities required in the production of one unit of an output may differ significantly from those involved in the production of another unit of the same type of output because of differences in case or client

[33] See "Purchaser Provider Systems" in Robinson, *Performance Budgeting*, for details on the failure of these systems.

characteristics. Output heterogeneity means that the cost of one unit of an output will differ from the cost of another unit of the same output for reasons which have nothing to do with efficiency. Other than in the quite special circumstance where the cost effects of heterogeneity average out over large volumes of production of the output concerned [34], significant output heterogeneity can undermine the foundations of an output-driven funding model. When a service is characterized by serious heterogeneity, its unit cost will be unpredictably variable, and it will therefore not be possible to base funding on unit cost. For example, the costs of treating accident victims in hospital emergency centers are too variable to permit them to be included within the DRG funding systems referred to above.

The fire brigade is a good example of a contingent capacity service. The output delivered by the brigade is putting out fires. But it would be impossible to fund the fire brigade on the basis of the numbers of fires it puts out. The fire brigade is like an insurance policy – government is not so much funding the output as the readiness of the brigade to provide the output quickly when and if it is needed. Deliberate spare capacity is built into such a service.

One of the reasons why attempts to apply the purchaser-provider system to the whole of government failed was that so many services provided by government ministries were unsuited to this method of funding. For these reasons, formula funding and purchaser-provider systems should be seen not as the basis for government-wide performance budgeting models, but rather as systems which can be selectively applied to selected government services – those which are relatively standardized, and produced in large volumes.

Key Readings

Robinson, M. (2007), "Purchaser-Provider Systems", in Robinson (ed) *Performance Budgeting*.

Smith, P. (2007), "Formula Funding and Performance Budgeting", in Robinson (ed) *Performance Budgeting*.

[34] Which is the case for many types of health treatments.

13. Targets and Performance Budgeting

Performance targets are an important managing-for-results tool. When targets are systematically linked in some way to budget funding, we can talk about a specific model of performance budgeting. The UK Public Service Agreement (PSA) system, in the form in which it operated between 1998 and 2007 under the Labour Government, is the most important example of this model of performance. Under the PSA system, several hundred high-level targets were set every three years as part of a "spending review" process in which multi-year ministry budgets are set. The spending review in fact constituted the core of the budget process – it was the context in which most ministry funding was decided. So the PSA system was one in which budgeting and target-setting were part and parcel of the same process.

Before considering the nature of the relationship between funding and targets, it is useful first to consider performance targets and some of the key issues which they raise.

A performance target is a quantitative goal with a timeline (usually explicit, but sometimes implicit) for the achievement of that goal. Targets may be set for outcomes, output (quantity, quality or efficiency), or even for activities or inputs. The distinction between indicators and targets is an important one. Targets are always based on specific performance indicators, which provide the yardstick for measuring target accomplishment. But indicators do not include quantitative objectives, nor timelines. Thus the percentage of HIV/AIDS infected persons in the population is a performance indicator. Cutting the rate of HIV/AIDS infection by 10 % over 5 years is, by contrast, a performance target.

It is also important not to confuse targets and objectives. An objective states the type of improvement sought without necessarily indicating the measure to be used, quantitative objectives or timelines. "Drastically reducing HIV/AIDS infection rates" is, for example, an objective rather than a target.

Performance Target: Some Examples

- Increase adult literacy from 60% to 70% by 2015 (an outcome target),
- Vaccinate the whole population against polio by 2012 (an output target),
- Ensure that all monthly accounting reports are completed within 15 days of the end of the financial year (an activity target, with implicit timeline of "immediately/this year"),
- Fill all vacant agricultural extension officer positions with suitably qualified persons during this financial year (an input target).

There is a widespread assumption that performance targets should be set for all programs and perhaps even for all performance indicators. However, it can be argued that selectiveness and gradualism in the setting of performance targets is preferable.

In the first place, target-setting requires a firm basis of good, timely and verified performance indicators. In leading OECD countries such as the UK, good performance indicator systems took decades to develop. Most countries – and in particular most low and middle income countries – do not have such systems. From this perspective, it seems to make sense in developing countries to focus initially on building a solid base of core performance indicators, and only at a later stage engage in extensive target-setting.

Secondly, setting appropriate and credible performance targets is quite difficult. Selecting which measures to turn into targets is not necessarily a straightforward matter, and the challenge of setting the quantitative targets which are neither too difficult nor too easy is quite considerable.

Thirdly, there are many who take the view that targets should only be set for performance indicators which are relatively controllable. The concern underlying this view is that setting targets for highly uncontrollable variables is more likely to de-motivate than to motivate agencies and their staff. If one takes this view, the most obvious implication is that targets should generally not be set for *outcomes*, and particularly not for high-level outcomes, because these tend to be greatly influenced by uncontrollable "external factors". From this perspective, setting targets for outputs will generally make more sense. [35] At the same time, however, outcomes are what matters most, and high-level outcomes are the outcomes which matter most of all. So many governments do in fact routinely set targets for outcomes over which they have quite limited control: e.g. rates of economic growth and unemployment, or rates of growth of tourist numbers.

A fourth and final point is that there is no point setting more targets that government can monitor and act on. In many countries, there is little follow-up of agency performance against target and, as a consequence, the targets seem not to have been taken very seriously. The UK experience under the PSA system was a major exception to this. There the government was very serious in monitoring and acting on performance relative to the PSA targets. Special attention was given to monitoring and following up on – including by means of managerial intervention when necessary – performance on the fifty most important PSA targets, via a "service delivery unit" reporting directly to the Prime Minister. The British experience also underlines that, if targets are to be taken seriously, it is crucial that they are target for

[35] With the qualification that targets may not be appropriate for outputs which are highly affected by heterogeneity.

indicators which really matter to the public and politicians and not, for example, targets relating to purely internal activities within ministries.

Because it makes little sense to set performance targets if one does not follow-up on performance against them, the more limited the capacity of central government to monitor and follow up on performance, the more selective central target-setting should be.

Gaming and Perverse Effects

The wisdom of a target-setting approach to improving public sector performance in contested by a school of opinion which regards targets as more likely to damage than improve performance. This school of opinion focuses on the danger that, to the extent that agencies and individuals within them are motivated to achieve targets, the targets may be met without a real improvement in performance because of:

- "Gaming" (manipulation or falsification of performance indicators, in this case of the indicators upon which targets are based) – for example, when a hospital falsifies its waiting time data in order to appear to have met a timeliness target which it has in fact failed to meet, and/or
- "Perverse effects" (deteriorating performance) arising from the use of targets based upon imperfect performance measures.

The problem of gaming requires careful verification and auditing of indicators, and sanctions for falsification. The problem of perverse effects, on the other hand, should in significant measure be minimized by careful design of, and the use of the right combination of, target measures. However, all performance measures are imperfect to a greater or lesser degree, so that even the most careful design of the set of targets cannot eliminate the potential for behavioral distortion.

Concerns about perverse effects are based in significant measure on theory. The most familiar theoretical point is there are some key dimensions of performance – such as quality – which are notoriously hard to measure, and tend therefore not to be captured in targets. The fear then is that what is not measured will be sacrificed to what is measure. In addition, theory points to the possibility that agents might pursue the easiest means of fulfilling their targets, with undesirable consequences. Theory is all very well, but empirical evidence on these questions is arguably more important. It is striking here that the empirical evidence of serious perverse effects under the UK PSA system – where the pressure to achieve targets was very considerable

– is very limited and the overall impression is that the critics have greatly exaggerated the problem. [36]

Judging on the basis of reported performance against targets, the UK PSA appears to have worked very well as an instrument for improving public sector performance, although the research on the subject has not been as in-depth as one might have wished for. The fact that targets were set in the budget process gave them real weight, as did the fact that performance against target was followed up by the Prime Minister and the finance minister, and was an important focus on the budget decision-making process. At the same time, it is clear that the connection between funding and targets was a very loose one. This is particularly the case because so many of the PSA targets were outcome targets, including some quite high-level outcome targets at that (e.g. targets for reductions in cancer rates and childhood obesity, over which the government clearly has quite limited control).

Key Readings

Audit Commission (UK) (2005), *Target Setting – A Practical Guide,* obtainable at http://www.idea.gov.uk/idk/aio/985665.

Smith, P. (2007), "Performance Budgeting in England: the Public Service Agreements", in Robinson (ed) *Performance Budgeting*.

Social Market Foundation (2005), *To the Point: a Blueprint for Good Targets*, London: SMF, obtainable at http://www.smf.co.uk/to-the-point-a-blueprint-for-good-targets.html.

House of Commons Select Committee of Public Administration (2003), *On Target? Government by Measurement*, obtainable at http://www.bercy.gouv.fr/lolf/downloads/1400_target.pdf.

Gay, O. (2005) *Public Service Agreements*, House of Commons Library, obtainable at http://main.hop.lbi.co.uk/documents/commons/lib/research/briefings/snpc-03826.pdf.

[36] A paper by Kelman and Friedman (2007) on the responses to a 4 hour target for treating patients in UK hospital accident and emergency rooms provides an antidote to fears about targets. Kelman and Friedman conclude that "that waiting-time performance improvement was dramatic and that dysfunctional responses, as far as we can tell, entirely absent." They add that "none of the hypotheses predicting effort substitution or gaming in connection with attaining this target has been confirmed", and that in fact dimensions of performance not captured in the targets appear to have improved. Bevan and Hood (2006) and Hood (2006) have undertaken important research which provides some limited examples of perverse effects, but which at the same time suggest that targets have worked well in raising performance (in Hood's words, the evidence "strongly suggests that targets made a marked difference in reported performance").

14. Public Financial Management Reform Foundations

Performance budgeting must be seen in the context of broader reform of public financial management (PFM) systems. Prior to any consideration of the adoption of performance budgeting is it, in particular, essential that the PFM system is able to deliver on two basic requirements. The first is that it is capable of supporting sound macro-fiscal outcomes. The second is that it broadly assures probity – that is, that money is spent only on public purposes duly authorized, and not widely diverted to private purposes by corrupt politicians and civil servants. If the PFM system and broader governance framework do not deliver these basic requirements, the initial focus should be on reforms in these areas, and performance budgeting should be deferred to the future.

In respect to sound macro-fiscal outcomes, fiscal sustainability is particularly important, because if budgetary policies are not sustainable it is likely that there will sooner or later be a major crisis in which public expenditure will have to be cut severely. Under the unstable circumstances which then arise, efficient and effectiveness management of expenditure inevitably suffers. It is only through the pursuit of sustainable fiscal policies that a sufficiently stable environment can be created in which public managers can successfully focus on results.

Although the principle causes of unsustainable fiscal policies are usually political, weaknesses in the PFM system can contribute significantly to the problem. Areas of PFM reform which can help greatly in promoting fiscal sustainability include:

- Expanding budget coverage so as to ensure that the budget is comprehensive – i.e. that it covers all expenditure (and revenue) which impact on deficits and debt.
- Ensuring in budget preparation that an aggregate expenditure limit compatible with fiscal sustainability is established and then respected. [37]
- Ensuring in budget execution that ministers and ministries respect the expenditure limits imposed upon them, and do not either overspend or accumulate payment arrears which will lead to future overspending.

[37] There has in this context been great interest in recent times in the development of more "top-down" budget processes. The reason for this is concern about the tendency of completely "bottom-up" budget processes – that is, processes in which spending ministries make unconstrained bids for extra resources during each annual budget process – to undermine respect for aggregate expenditure limits. In its extreme form, top-down budgeting counters this via unilateral central determination of the budgetary spending authorization of each spending ministry. In practice, a completely top-down process is neither feasible (other than in the very short term) nor desirable. However, fiscal sustainability may be strengthened by making the budget process *more* top-down in important ways, including – when medium-term budgeting is well-developed – the entirely top-down setting of envelopes *for existing programs* (as distinct from new expenditure initiatives).

Assuring probity is, of course, a matter of degree. No country is entirely free of corruption in the use of public monies. The question is therefore whether the system is capable of ensuring an acceptable minimum level of probity. Particularly important here is the quality of the parliamentary and other controls which are limit executive discretion in choosing the purposes for which public monies may be spent – that is, which are designed to ensure that monies can only be authorized for public purposes. The legal framework of budget approval is crucial here. The other aspect is safeguards during budget execution to ensure that money is spent in the manner intended. This requires sound expenditure control processes, as well as good internal and external audit.

For both sustainability and probity, a strong system of commitment control is important. Commitment is the stage where an obligation to make a future payment is entered into, such as when a contract to purchase a good or service is signed with a supplier. Once a commitment is legally entered into, it becomes essentially impossible to avoid the subsequent payment. A weakness in the PFM systems of many countries is the absence of sufficiently strong discipline at the commitment stage, with the result that commitments are entered into which are either or both excessive and for purposes inconsistent with the budget authorization. Sound commitment control should therefore be regarded as one of the foundational elements of good PFM. However, approaches to commitment control vary greatly. In countries where civil service discipline in weakest, the MoF will need to directly control commitments itself – that is, to require that no commitments be entered into without its approval. In more advanced systems, commitment control is decentralized to ministries and managers within those ministries, backed by sanctions which are applied if excessive or inappropriate commitments are entered into.

It is therefore important, before deciding to adopt a performance budgeting system, to take stock of the overall PFM framework and to determine whether there are weaknesses in these fundamental areas which should be addressed first. Particularly useful in this context is the PEFA framework – a standardized international diagnostic framework for public financial management systems. [38]

There are a number of areas of PFM and related systems which are *co-requisites* (rather than prerequisites) of performance budgeting. This means that while it is not essential that reforms in these areas are in place *prior* to the start of the performance budgeting implementation process, these reforms must *accompany* performance budgeting if it is to succeed. The importance of such accompanying reforms is demonstrated by the unfortunate examples of many countries which have implemented key elements of performance budgeting (in particular, program classification of the budget and the development of program performance

[38] See the PEFA website (www.pefa.org) for more details.

information) but found that this in itself did not change budget outcomes: in other words, expenditure prioritization did not improve and there was little discernable impact on efficiency and effectiveness.

Other sections of this manual discuss two critical co-requisites: namely, improved expenditure prioritization processes in budget preparation, and civil service performance management reform to give civil servants the incentives and freedom to focus more on results.

The other area of PFM reform which is critical is the reduction of expenditure inflexibilities. Expenditure inflexibilities are obstacles to the reallocation of resources. Because government-wide performance budgeting is critically concerned with the reallocation of resources from low-priority and ineffective programs to high-priority ones, it will never succeed if expenditure inflexibilities make such reallocation impossible or very slow and difficult. Amongst the inflexibilities which can be of particular concern are:

- Extra-budgetary fund arrangements which bind certain revenues to be used only in designated program areas.
- Civil service employment inflexibilities which, in some countries, make it impossible to shift or terminate the employment of civil servants who work on programs which the government wishes to eliminate or cut back.

Creating more flexible civil service employment conditions is a task which must be approached in different ways in different countries, depending upon their governance arrangements and traditions. In certain OECD countries, traditional civil service guarantees of job security have been largely abolished and the civil service made to operate under general labor law. Such a radical approach may not, however, be appropriate for countries where there is great concern about the scope for abuse of civil service employment flexibility for political or nepotistic reasons. However, even if much more limited reform is adopted, it is essential at least to make it possible to re-assign workers and to make use of strategies such as voluntary departure packages to downsize when necessary.

Key Readings

Diamond, J. (2007), “Challenges to Implementation” in Robinson (ed.), *Performance Budgeting.*

15. Expenditure Prioritization and Performance Budgeting

Improved expenditure prioritization and increased performance pressure on ministries and agencies are the two channels by which government-wide performance budgeting aims to improve public sector performance. However, as noted previously, merely producing information on the benefits and costs of programs does not ensure that this information will be used to improve prioritization and hold ministries to account for performance. There need also to be formal routines for the reconsideration of spending priorities integrated into the budget process, and these routines need to be designed so as to make maximum use of available information on program performance.

Spending Review

The key point of contact between performance budgeting and expenditure prioritization processes during budget preparation is *spending review*. Spending review refers to the systematic scrutiny of existing expenditure to identify, in particular, options for cuts. Spending review draws on evaluation (see Section 6). More specifically, it draws on both program evaluations (the review of specific services provided by government) and efficiency reviews (which focus on reducing the cost of delivering services). However, spending review also goes beyond evaluation to include systematic *priority* analysis – in other words, the systematic identification of programs or elements of programs which could be cut *because they are low priority*. This is a completely different matter from the evaluation of ineffectiveness or inefficiency. A program might be highly effective and efficient, but still be very low priority because the outcomes which it aims to achieve are not very important to the community, or are not rated as such by the government of the day.

Without spending review, the risk is that programs which are ineffective, low-priority or which have outlived their usefulness will continue to command public resources. It is in the process of spending review that performance indicators and evaluations can be systematically employed, in conjunction with the measure of program costs which a program budgeting system generates.

Spending review is an area where the budgeting systems of many countries are weak. In such countries, the budget process is overwhelmingly about new spending, and ongoing expenditure is not generally speaking seriously scrutinized. "Incrementalism" is a term coined by budgeting

writers in the 1970s to describe this tendency of budgeting to take spending on existing programs for granted.

Spending review is critical to good aggregate fiscal outcomes and to the capacity of the government to respond to new spending needs. If *substantial* room is to be created for important new spending initiatives, it will almost always be necessary to cut existing spending. This is important also for aggregate fiscal discipline, because if such cuts are not identified, the danger is that new spending will simply be added on to the budget, pushing up aggregate spending at a higher rate than is consistent with keeping the budget deficit at sustainable levels.

Good spending review also puts increased pressure to perform on spending ministries, because it greatly increases the probability that poorly-performing programs or areas of inefficiency will be identified by the center and result in either budget cuts or sanctions being applied to ministry management. Ministries which wish to protect their budgets will as a result be motivated to lift their performance. In this context, spending review should also be linked with processes for management improvement and program re-design. This is because, if a program is identified as ineffective, it will not necessarily follow that its funding should be cut: a change in program design or management may be more appropriate.

Spending review should be integrated with the budget process. In most countries, this will mean that some spending review is undertaken every year as part of the annual budget process. In countries (such as the UK) where fixed medium-term expenditure ceilings are set for spending ministries, spending review is a process which is carried out only every three or four years (see Section 18). In most of the following discussion, it is assumed that spending review is an annual process.

Advocates of *zero-base budgeting* favor the comprehensive review of all expenditure every year, or at least at regular intervals. Experience shows, however, that this is not practicable – the analytic task is simply too great. Spending review needs instead to be selective and strategic in its focus. The question then arises as to what approach should be taken to selecting programs for review. One approach is the discretionary targeting approach, in which spending review is focused on programs which officials and ministers believe to be most likely to yield savings. The alternative is a regular program review cycle in which all programs are reviewed over a regular multi-year cycle of, say, five years. Canada is one example of a country with such a system (see Box), and the U.S. Program Assessment Rating Tool until recently involved a

similar five-year cycle in which 20 percent of US federal programs were subject to review each year.

Spending Review Cycle: The Canadian Example [39]

The Government of Canada introduced a new expenditure management system in 2007 as part of an ongoing commitment to better manage government spending. This system aims to ensure value for money for all government spending. A key pillar of this system is the ongoing assessment of all direct program spending, or strategic reviews. Through the Strategic Review process, the Government systematically assesses the relevance and performance of every program on a cyclical basis. The Strategic Review process requires government organizations to review 100 percent of their direct program spending and the operating costs of their major statutory programs on a cyclical basis. From this review, the organization identifies five percent of spending for reallocation from its lower performing, lower priority programs. This system is combined with a system of targeted strategic reviews based on the selection by ministers and the Treasury Board Secretariat of selected programs for review.

The scope of expenditure covered by the spending review should be the whole of general government, and should certainly not be limited to "discretionary" expenditure. It is, in particular, crucially important to include social entitlements and other mandatory spending.

Spending review involves three levels of activity. At the lowest level are reviews of individual program, efficiency and other topics. The next level is the preparation, on the basis of these reviews and other analyses, of coordinated advice to the political leadership identifying the best options for expenditure cuts. The final level is decision-making about which cuts to make, which is the role of the political leadership.

Processes for the political leadership to make prioritization decisions should therefore be well-integrated into the budget process. The nature of the most appropriate mechanisms for political decisions about priorities will vary greatly between countries. In some countries, a process involving a Cabinet committee may be the best approach. In other countries, the political system may dictate a more centralized approach, focused on the president and/or prime minister in conjunction with the finance minister. The role of the parliament is discussed in Section 17, reflecting the fact that there is great variation internationally in the extent of the power which parliament exercises over the allocation of resources in the budget.

[39] Source: Treasury Board Secretariat, http://www.tbs-sct.gc.ca/sr-es/faq-eng.asp#q1

Cabinet Committees and Expenditure Reallocation: The Australian Example

The use of a cabinet committee as the instrument for key decisions about expenditure reallocations has been a distinctive feature of the budget process of the Australian national government since the 1980s. Throughout this entire period, there has been a cabinet "Expenditure Review Committee" (ERC) charged with examining major new spending proposals put forward by spending ministries, and also with discussing major cuts which could be made to finance new spending or reduce the deficit. The ERC is comprised of the Finance Minister and the Treasurer (who share responsibility for the budget in the Australian system), together with the Prime Minister and one or two powerful spending ministers. A key function of the ERC is to strengthen the hand of the guardians of spending discipline within the Cabinet, thus countering the tendency of spending ministers to form coalitions to push spending up. The analytic support of finance ministry staff has been crucial to the success of this mechanism. MoF officials provide ERC members with so-called "Green briefs" for each spending ministry, which not only critically evaluate the ministry's new spending proposals, but identify savings options and alternative strategies. They are also present at meetings of the ERC to be able to advise ministers on the spot, and to act as a counterweight to spending ministries. The Labor Government elected in 2007 modified the ERC process to some extent by instituting a Strategic Budget Committee (SBC), comprised of the Prime Minister, the Deputy Prime Minister, the Treasurer and the Minister for Finance, which met at the very beginning of the budget process and assumed the function of overviewing spending ministers' new spending proposals. The ERC was, however, retained, and pursues its work in the light of the parameters established by the SBC. The Government also reverted in 2007 to practice previously employed in times of significant fiscal consolidation – namely, of holding meetings of the ERC throughout the year, rather than only during specific stages of the budget process.

The provision of advice to the political leadership on options for spending cuts, arising from spending review, should – depending on the institutional structure of the country – be the function of the MoF, potentially in collaboration with the president's or prime minister's office. This advice needs to be based on policy analysis as well as financial considerations. Only civil servants can carry out this function on an ongoing basis—external advisers do not, in general, have a sufficiently in-depth knowledge of government, and are not able to provide the continuity required. Nevertheless, external advice has its place, and in carrying out the spending review function, the MoF may draw in part on program and efficiency reviews carried out by external consultants or commissions.

Spending review should also in part be carried out by the spending ministries themselves, which should be encouraged to develop their own internal program evaluation. But spending review cannot be left to the spending ministries alone. The MoF must lead it, and in doing so should not allow itself to become dependent on program evaluations carried out under the exclusive control of spending ministries.

Expenditure Prioritization Processes More Generally

As noted above, spending review is the crucial point where performance budgeting interacts with expenditure prioritization processes. To work well, however, the expenditure prioritization process requires other key elements.

Good processes for determining priorities for *new spending* are crucial. Two elements are particularly important here. The first is the existence of an effective *strategic phase* in the budget preparation process – that is, of a stage early in the budget preparation process where the political leadership sets the priorities which will guide the process. A *budget strategy paper*, presented to cabinet of ministers for consideration at the start of the budget preparation process, can be one good vehicle to facilitate the determination of government priorities during the strategic phase.

The second key element of importance in relation to the prioritization of new spending is rigorous processes for central review of these proposals. Particularly useful in this context can be:

- Strict requirements about the manner in which proposed new spending initiatives are presented: what type of information is to be provided about the proposal (e.g. a clear statement of the objective), a requirement for medium-term estimates of the cost of the initiative, minimum amounts of notice to be given, a requirement of circulation of the proposal to a specified list of central agencies who will analyze and comment upon the proposal.
- A requirement that the cost estimates for each new initiative be agreed by the MoF, so as to prevent spending ministries deliberately underestimating the costs of their proposals.

Proposals for new spending should to the maximum degree be integrated with the budget process. In some countries, a large portion of proposals for new spending by spending ministries are put forward and decided by the government outside the budget preparation process. As a consequence, during budget preparation, the focus is not so much on making decisions about new spending options as on trying to fit within the budgetary resource limits the spending required by new laws passed prior to the budget. An obvious danger under such circumstances in that the new spending initiatives adopted outside the budget process will cost more than the government can afford. The other adverse consequence is weaker expenditure prioritization. This is because, with new spending proposals being put forward at any time

during the year, they tend to be considered in isolation rather than being compared with other possible new spending options in the way that would happen if new spending proposals were in general considered jointly during the budget preparation process.

Countries are sometimes advised to set firm ministry (or sectoral) budget ceilings right at the start of the budget preparation process, in an entirely "top down" process prior to any budget bids or other bottom-up input from spending ministries. This advice is based on the proposition that a completely "bottom up" budget process, in which ministries are able to bid for whatever they like with no indication of resource constraints, gives ministries no incentives to identify savings, and inevitably results in bids so far in excess of available resources that they are impossibly difficult for the MoF to handle. [40]

The concern about purely bottom-up budget processes is a valid one. There is, however, a real danger that, unless the ceilings are formulated very carefully, setting early and firm ministry ceilings will aggravate allocative rigidity, greatly limiting the possibilities for improved expenditure prioritization. Neither ministers nor central agencies such as the MoF have the detailed knowledge of the program areas to unilaterally determine where new spending should be focused. Moreover, the amounts of new money which ministries or sector should be given for new spending can only properly be decided after the examination of concrete program proposals. It makes no sense to say something like "health is a top priority sector, so we'll increase its funding by 15 percent without considering what concretely the health ministry will spend the extra money on".

Few OECD countries in fact set ministry ceilings before considering bottom-up proposals for new policy. Countries like Canada and Australia manage to avoid the dangers of a purely bottom-up budget process while nevertheless retaining considerable scope for bottom-up ministry proposals for new spending initiatives. They do this by requiring spending ministries to submit major new policy proposals separately from their core budget proposals, with the latter based on the continuation of pre-existing policies and programs. They then subject the new spending proposals to very rigorous and systematic central scrutiny. Chile did the same thing some years ago with its well-known "Bidding Fund".

Sweden is sometimes seen as a model for purely top-down ministry ceiling-setting. In theory the ministry ceilings are all set at a cabinet retreat one month into the budget preparation process, prior to any bottom-up input. However, in reality, discussions on new policy initiatives which impact on the ceilings continue for months afterwards.

[40] See Schiavo-Campo & Tommasi (1999) and Potter & Diamond (1999).

Good prioritization requires the right blend of top-down and bottom-up elements in the budget preparation process. It certainly will not be facilitated by an entirely bottom-up process. Equally, however, a totally top-down process is neither desirable nor workable, other than temporarily under extreme circumstances (e.g. an acute fiscal crisis requiring quick and drastic expenditure cuts). Defining precisely the right balance between top-down and bottom-up processes during budget preparation arguably depends very much on country circumstances.

Planning and Prioritization

The prioritization of expenditure is, of course, is the key function of *planning.* Planning systems are, nevertheless, often weak at prioritization. This is because they are often better at identifying purposes for which the government should spend more, than at identifying where spending could be cut back to make room for these new spending priorities.

Moreover, where the planning process is institutionally separated from the budget process, budget decision-makers may not take the priorities identified in the plan seriously. The problem tends to be particularly serious in countries which think of planning and budgeting as sequential processes. That is, the plan is prepared first and is where all the priorities are formulated. The budget is then supposed to merely give financial expression to the plan's priorities. Often, this leads to the plan being formulated out of the context of budgetary constraints, with the result that the plan doesn't really prioritize in the sense of making the hard resource allocation choices.

There are two key lessons to be learnt from this: firstly, that planning and budgeting must be fully integrated. Expressed differently, planning should take place as an integral part of the budget process, rather than being something which is undertaken only prior to the preparation of the budget. Secondly, good planning has to be as much about identifying cuts as about selecting areas for new spending – that is, it has to be designed to prioritize expenditure, not just to dream up "blue skies" ideas for new spending.

Expenditure Prioritization at the Sector and Ministry Levels

One influential approach to the prioritization of expenditure has been the so-called "sector-wide" approach. In this approach, much of the task of expenditure prioritization is delegated to sector groups of ministries. The centre of government would initially determine ceilings for sectors such as, say, agriculture, and then all of the ministries with a role in that sector would

get together to decide the optimal allocation of that ceiling between sector programs (after a discussion of sector objectives and priorities). [41] This system appears to have been inspired by the Canadian Policy and Expenditure Management system (PEMS) which attracted considerable favorable attention in the early 1980s. [42]

This system has certain advantages. However, Canadian experience suggests that multi-ministry committees are not necessarily a good forum for spending review. Moreover, it can also be argued that the sector-wide approach delegates too much of the prioritization task. The center is supposed to decide the sectoral allocation, but it is hard to decide how much money sectors should get except in the context of specific major measures or savings which are proposed.

The sector-wide approach is correct at least in its recognition of the need for significant decentralization of prioritization decisions. However, some would suggest that this may be more appropriately done at a ministry rather than sectoral level.

Given that it is neither desirable nor feasible to centralize all prioritization decision, a key question is how to create incentives and pressures for spending ministries to allocate the resources at their discretion in the most effective manner. Two important mechanisms can be mentioned: (1) a system whereby ministries are expected to fund minor new spending initiatives through internally-identified savings, and are therefore given an incentive to find such savings, and (2) the pressure imposed by good central spending review (as discussed earlier).

A good expenditure prioritization process, integrated fully in budget preparation, is critical to the success of a government-wide performance budgeting system. Without good prioritization processes, performance information may not be actually used in budget decision-making, and the results may be that all the effort of developing indicators, evaluations and program costings may end up having little impact on the allocation of resources.

Key Readings

[41] An outline of this approach in the MTEF context can be found in the World Bank (1998) *Public Expenditure Management Handbook*, pp. 47-8. In this version, the sectors also undertake a range of other related performance planning and management roles, including the explicit definition of sector objectives, performance indicators and targets.

[42] See D. Good (2008) *The Politics of Public Money* (University of Toronto Press).

Kim, J. M. and C-K. Park (2006), “Top-down Budgeting as a Tool for Central Resource Management”, OECD Journal on Budgeting, 6(1): 93-94.

Ljungman, G. (2009), *Top-Down Budgeting – An Instrument to Strengthen Budget Management*, IMF Working Paper, obtainable at http://www.imf.org/external/pubs/ft/wp/2009/wp09243.pdf.

OECD (2005), *Reallocation: the Role of Budget Institutions,* OECD: Paris, can be ordered at http://www.oecd.org/document/38/0,3343,en_2649_34119_33701862_1_1_1_1,00.html.

Robinson, M. (2011), “Keeping the Lid on Aggregate Expenditure during Budget Preparation: Enforcing Aggregate Expenditure Ceilings while Preserving Allocative Flexibility”, in proceedings of Banca d’Italia Fiscal Policy Workshop 2001, available at http://www.bancaditalia.it/studiricerche/convegni/atti.

Robinson, M. (2010), *Ministry Ceilings under Top-Down Budgeting,* at http://blog.pfmresults.com/wordpress/?p=10.

Robinson, M. (2010), *Keeping the Lid on Aggregate Spending*, at http://blog.pfmresults.com/wordpress/?p=93.

16. Accrual Accounting and Performance Budgeting

This section focuses on the relationship between accrual accounting and performance budgeting. There are some who say that accrual accounting, and perhaps also accrual budgeting, is essential for performance budgeting. We consider in this section whether this is correct. The conclusion is that it depends on the form of performance budgeting. There are certain forms of performance budgeting for which accrual accounting is indeed essential. But this is not true of most forms of performance budgeting. It is, in particular, not true for program budgeting. Although program budgeting may arguably be more effective in an accrual environment, it will work perfectly well under traditional "cash" (or "commitment") budgeting and accounting.

The section starts with an explanation of accrual accounting (AA) and accrual budgeting, before moving on to consider their relevance for performance budgeting.

What are Accrual Accounting and Accrual Budgeting?

AA is a financial reporting system. In other words, it is a specific methodology by which organizations report their financial transactions and position. AA is the standard accounting methodology of the private sector, where it arose primarily as a method for properly measuring enterprise profit. AA may be defined as a methodology by which entities may properly measure their costs of production and, where relevant, revenue earned.

In a private sector context, it is only through the proper measurement of costs of production and revenue that profit may be correctly measured. The correct measurement of revenue earned and profit or less is also relevant to government business enterprises (GBEs), because they also earn revenue from selling products to consumers, and make a profit or loss on their operations which needs to be measured. It is precisely for this reason that GBEs in most countries have therefore long employed AA.

By contrast, the concepts of revenue earned and profit are largely irrelevant to *government* bodies – that is, to ministries and other organizations which are primarily dependent on tax revenues for their financing. For them, the potential benefits of AA arise largely from its use to measure their costs of production – that is, the costs of the outputs which they deliver. For this

reason, in explaining AA, we focus firstly in on the accrual treatment of costs, rather than of revenues.

Traditionally, governments around the world have used so-called cash accounting and budgeting (coupled, in some countries, with so-called "commitment" accounting and budgeting) in their budget sectors. Only relatively recently have some governments moved to replace cash accounting with AA. Under cash accounting, the principal focus of financial reporting is upon payments and receipts of money during the financial year. At the ministry level, the main form of financial reporting is the reporting of ministry expenditure, with expenditure defined as payments to external parties [43]. Under cash *budgeting*, the spending budgets given to each government ministry are also expressed in cash expenditure terms. In other words, each ministry is given an annual budget formulated as a limit on the payments it is permitted to make during the year. [44]

The easiest way of understanding AA is to contrast it with cash accounting. In other words, we can best understand the nature of accrual accounting by explaining why it is that cash accounting is inappropriate as a basis for measuring the costs of production, and how it is that AA overcomes the weaknesses of cash accounting in this respect.

If a ministry wishes to work out the costs of production of the services it has delivered to external clients this year (i.e. of its outputs), it will need to take into account the costs of all resources used to produce those services. If it tries to do this by looking at this year's cash *expenditure* – that is, on the basis of its cash accounting – it will face a number of difficulties. One is that some of the resources used to produce outputs this year will have been paid for in previous years (e.g. buildings, equipment and supplies purchased in the past). Equally, some resources used this year may not be paid for until sometime in the future (e.g. bills for supplies delivered and used at the end of this financial year, which are not payable until the beginning of

[43] More precisely, non-repayable payments: in other words, payments other than repayments of public debt and loans made by government to external parties.

[44] In countries with commitment budgeting, there would also be a limit on the value of commitments to make future payments which the ministry could make during the year. A "commitment" refers to a legal commitment to make a payment, whether at the time the commitment is made or at some subsequent point of time (potentially even some years in the future, in the case for example of multi-year contracts with, say, construction companies for major public works). In a commitments budget, quantitative limits are imposed on the quantum of new commitments which may be entered into during the financial year in question.

the next financial year). A focus on cash expenditure will overlook these 'non-cash' costs, and to that extent will understate the costs of production.

The problem, moreover, cuts the other way as well. A measure of this year's cash expenditure will include all capital expenditure undertaken during the year. To treat this year's expenditure as a measure of this year's costs of production therefore involves an implicit assumption that all *capital* expenditure which takes place this year contributes to production in this year and this year alone. By definition, however, capital expenditure is expenditure on assets which contribute to production over a number of years. This means that to count all of this year's capital expenditure as part of this year's costs of production is to *overstate* costs.

AA solves the above problems by replacing the cash expenditure measure with a measure of the cost of the resources used in production, known as *expenses*. The expenses of an entity in a given financial year measure the costs of all of the resources the entity uses to produce outputs in that financial year, irrespective of when those resources where actually paid for. The difference between the expenses concept and the expenditure concept therefore concerns the financial year in which payments for inputs used in the production process are recorded. *Expenditure* measures when the inputs are paid for, whereas *expenses* measures when they are used in the production process.

Take the example of capital expenditure. Rather than recording all capital expenditure as a cost of production in the year the expenditure takes place, AA treats a portion of the capital expenditure as a cost of operation (an expense, in accrual language) in each of the years in which the assets concerned contribute to production. Roughly speaking, if an asset's life is, say, ten years, AA might count one-tenth of the price paid for the asset as a cost of production in each of the ten years of its life. [45] This is what is known as "depreciation".

The difference between the cash accounting concept of expenditure and the AA concept of expenses does not arise only in relation to the treatment of capital. Another example relates to bills (accounts) presented to a department by its suppliers. Under cash accounting, such bills are recognized in the department's financial reports only in the financial year when the

[45] This simplifies the concept of depreciation greatly. In practice, the accrual treatment of capital expenditure is considerably more complicated, when one takes into account things such as inflation and the possibility that the asset has some scrap value at the end of its life, let alone the more complicated asset valuation and depreciation methodologies accountants have developed. However, the idea of the accrual treatment of capital as cost allocation over time illustrates the essence of the accrual concept of expenses as applied to capital assets.

expenditure actually takes place (i.e. when the bill is paid). By contrast, under AA bills are recognized as an expense in the financial year when the relevant supplies are delivered and the bill therefore become payable (i.e. when the department incurs the liability to pay the suppliers). This distinction can make a difference because, as noted above, payment for some supplies may not take place until the financial year after the relevant they have been received and used in production.

Civil service superannuation – pension costs – offers another example of the accrual treatment of costs. Most governments operate their own superannuation schemes. In other words, they promise retirement benefits to their civil servants. This makes the retirement benefits entitlements which civil servants build up during their careers obligations which the Government itself must meet. Consequently, from the Government's point of view, these entitlements represent a major financial liability. Moreover, they are an important part of the cost of delivery of government services, because during each year a civil servant is employed they build up additional superannuation entitlements. The question therefore arises as to how should one measure the superannuation component of the Government's costs of production in any particular year. Cash accounting is no good for this task, because it is not possible to measure the annual *cost* of superannuation by looking at annual superannuation *expenditure*. To see why not, consider the position of an individual public servant, Jim. Each year during Jim's career, he builds up a growing entitlement to superannuation benefits. However, the actual payment of those benefits (i.e. the expenditure) takes place only after the conclusion of Jim's career. If one employed cash accounting, we would have the nonsense result that:

- No superannuation 'cost' would be recorded for Jim during the years when he was working and providing services to the public,
- There would be a substantial 'cost' recorded only when Jim was retired and manifestly making no contribution at all to production.

What AA does is to recognize the fact that during each year of Jim's career, there is a superannuation cost attached to his employment. This cost is, of course, a non-cash expense (i.e. it involves no payment at that time), but this makes it no less real. How is the superannuation cost of public service employment measured under AA? Essentially, the idea is that instead of measuring the expenditure on superannuation benefits, the government as a

whole measures its annual superannuation costs by counting the gross increase in the total amount of superannuation entitlements which government 'owes' its employees. [46]

It will thus be clear that only by using AA can public sector agencies accurately measure the cost of producing the outputs which they deliver to the public. Of course, what AA generates is a total figure which measures the aggregate cost of all of the outputs produced by an organization in a particular financial year. AA per se does not provide information on the cost of producing a *specific* output or group of outputs, nor does it provide information on the *unit cost* of outputs. To obtain such information requires accrual accounting measures of expenses to be brought together with output information. It is a function of so-called *management accounting* to split the total accrual expenses of an agency between its various programs (or, at a more detailed level, between specific types of services).

The discussion to this point has focused on accrual accounting. Accrual *budgeting* is the use of accrual accounting as the basis for budget allocations – more precisely, the use of accrual concepts to specify budgetary *control totals*. [47] As discussed in the section on program appropriations, control totals are the quantitative spending limits imposed on spending ministries, whether as appropriations in the annual budget law or by administrative directive (e.g. from the minister/MoF). Accrual budgeting must therefore not be confused with accrual accounting, which refers only to the recording and reporting of financial operations of government in accrual terms. Governments may *budget* in cash terms while *accounting and reporting* in accrual terms [48].

What does using accrual concepts to specify budgetary control totals mean? The most basic type of accrual control total is an *expenses control total*—that is, a centrally-imposed quantitative limit on the expenses each spending ministry is permitted to incur within the budget year. [49] An expenses control total means that current year expenses are counted as the

[46] "Gross" means here the increase in superannuation entitlements ignoring the quantum of the entitlements of already-retired civil servants which the government pays off during the year.
[47] More formally, accrual budgeting has been defined (Robinson, 2009) as *the specification of budgetary expenditure authorizations and revenue estimates in terms of accrual accounting measures*.
[48] Note here that accrual accounting reports not only on expenses etc, but also on cash flows, and may therefore be said to incorporate cash accounting.
[49] As explained later, it is not essential that an expenses control total cover all categories of the ministry's expenses. The imposition of an expenses control total covering most categories of each spending ministry's expenses should, nevertheless, be regarded as a core feature of any accrual budgeting system.

use of the budgetary funding, irrespective of the timing of any associated cash payment. For example, the following would be counted against the expenses control total:

- Bills payable: amounts which are owed for goods and service delivered and used during the financial year, even though payment may not be made until the next financial year.
- The additional entitlements for future pension payments which civil servants accumulate during the financial year.
- Stocks (of supplies etc) actually used in the production process during the year, irrespective of when those stocks were purchased.
- Depreciation of the ministry's fixed assets.

The Complexity of Accrual Accounting and Budgeting

The core case for AA in respect to performance budgeting rests on the proposition that, the better the measure of the cost of outputs, the more useful this will be for performance budgeting. As discussed below, this is – broadly speaking -- true. However, a judgment about whether AA should accompany performance budgeting cannot be made without considering its costs as well as benefits. In this context, it is important to recognize that AA has significant downsides, the most important of which is that it is more complex and, as a result, costs more in skilled labor and accounting system terms to operate. These costs and skilled labor demands make the introduction of AA particularly challenging in low-income (and even middle-income) countries. The question which each country must therefore ask is whether the benefits of introducing AA into government outweigh the costs. In respect to performance budgeting, it is not sufficient merely to demonstrate that program budgeting or some other performance budgeting system can in theory work better in the presence of AA. What needs to be demonstrated is that the improvement is sufficiently great (taken in conjunction with other benefits of AA) to justify the costs and complexities of the move to accruals. It also needs to be borne in mind that both performance budgeting and AA and demanding reforms in their own right. It may not necessarily make sense to attempt to introduce them both at the same time.

This is even truer of accrual *budgeting*, which is considerably more complex that AA alone. The additional complexity not only means that an accrual budgeting system is highly demanding of skilled human resources (accountants and others) and supporting IT systems, but that it may weaken expenditure control in countries where accrual concepts are not well understood by managers (see Robinson, 2009).

Accrual Accounting and Program Budgeting

As repeatedly emphasized, the primary objective of program budgeting is improved expenditure prioritization. Program budgeting improves expenditure prioritization by making it possible for budget decision-makers to compare (usually in an informal way) the costs and benefits of alternative programs.

The discussion above makes it clear that cash accounting is imperfect for this task. Because cash accounting does not properly measure the costs of service delivery, it can distort program cost comparisons by making some programs look less costly than they really are – and less costly relative to other programs than they really are. Suppose, for example, budget decision-makers wish to compare a health treatment program and a preventative health program. Health treatment is very capital-intensive because a great deal of expensive equipment and buildings are used. Health prevention relies on information campaigns, and makes little used of costly infrastructure. If one compares these two programs on a cash accounting basis – that is, by looking at the cash expenditure which each undertakes – during a period when there is not much new investment in health treatment facilities underway (i.e. not many new hospital being built etc), then the comparison will be distorted. The health treatment programs costs will look artificially low, while this will not be the case for the preventative health program.

Other distortions can arise when making comparisons between programs as a consequence of using cash accounting. For example, if we are comparing a program which is delivered primarily by civil service staff to another which relies largely on contracted private-sector companies and their employees, the considerable deferred employment cost component of civil service employment may make the former program look relatively cheaper than it actually is.

It is clear, then, that it is in principle better to base a program budgeting system on AA. AA measures of program costs facilitate much fairer comparisons of the costs of alternative programs. Accrual budgeting would be even better, because then the full cost of programs would be charged to ministry programs, giving them a powerful incentive to base their prioritization decisions on the real cost of programs, and not just on the level of cash expenditure that they require in any given year.

The Link between Accruals and Performance Budgeting

"One of the key elements of the new resource based [i.e. accruals] approach is that it requires ministries to undertake more accurate costing of activities, with expenses and income allocated to each of a ministry's objectives. This will assist the Government in ensuring that resources are allocated to priority services in line with the Government's objectives." (HM Treasury, 2002)

"The introduction of cost [i.e. accruals] principles into government budget planning and accounting has to do with improving government effectiveness through visibility, transparency, and cost-conscious behavior. … Transparency of costs provides politicians and administrative leadership a better basis for prioritizing use of resources." (Danish Ministry of Finance, 2006)

However, even though cash accounting creates some distortion in measures of comparative program costs, it cannot be said that these distortions are on the whole so serious as to prevent program budgeting from achieving its basic objective of improved expenditure prioritization. Program budgeting has worked well in some countries with cash (or cash and commitment) accounting and budgeting systems. Moreover, amongst the factors which determine whether program budgeting works or not, the precision of the accounting system in measuring program costs is in most countries a much less important consideration that whether the budget process is designed to facilitate good prioritization decisions (see the Section 15). So in terms of making program budgeting work well, the introduction of AA is usually not the most important priority.

For developing countries in particular this is good news. They should not feel that they must combine the introduction of performance budgeting with a move to AA, let alone accrual budgeting. They should bear in mind that most program budgeting systems – including most recently the well-designed system introduced in France in 2001-2006 – have operated in a traditional cash or cash/commitments accounting and budgeting framework.

Accrual Accounting, Purchaser-Provider and Formula Funding Systems

By contrast to program budgeting, a purchaser-provider (PP) system absolutely requires AA. As discussed Section 12, PP requires that government agencies operate like businesses, with a focus on their financial bottom line (profit or loss). There is only one way of measuring profit and loss – through the use of AA. The use of AA is essential for this purpose not only because it

is necessary to measure production costs accurately, but also because the accurate measure of revenue earned is absolutely imperative.

The AA approach to the measurement of revenue applies exactly the same principle as is applied in the measurement of expenses. That is, accrual revenue is recorded in the year it is actually earned, as opposed to the year when the cash happens to be received. The difference between cash revenue and accrual revenue is then, once again, about the financial year in which the revenue is recorded. Suppose, for example, that a business received a large payment this year for delivery of services to be made over the following, say, three years. Cash accounting would record this all as revenue this year. But this would be absurd, as the revenue concerned will in reality only be earned progressively over the next three years as and when the product is delivered. Reflecting this, AA would recognize the revenue in stages over those three years.

Similarly, if an enterprise sells an asset, cash accounting would treat this as revenue. By contrast, accrual accounting would see that the cash received is not really revenue earned, because it was received only by sacrificing the asset concerned. From the accrual perspective, to treat such receipts as revenue would be rather like someone who sells their car treating the payment they receive as if it were equivalent to salary.

Hence the most important of the suite of financial statements generated in an AA context: the operating statement, commonly referred to in the private sector as the Profit and Loss Statement (P&L). The operating statement reports total revenue and total expenses in the financial year concerned (identifying the key elements of both). The 'bottom line' of the operating statement is the operating result, which is total revenue minus total expenses. The operating result is, of course, a measure of accounting profits.

Any hospital or other agency which is funded via a PP mechanism rather than traditional budget funding needs to measure its operating result, just like a private business. Under the PP system, the agency concerned needs to be sure that it is not making a loss, and AA provides the only means of assessing this. It is precisely for this reason that when Australia and New Zealand attempted (unsuccessfully) to apply the PP principle to the whole of the government budget, they coupled it with a move to accrual budgeting – hence the term “accrual output budgeting”.

In respect to formula funding (FF), the story is a little more complex. As explained in the relevant section, some FF systems are based on costs, and it is in relation to these that AA has a potential contribution to make. Cost-based FF is generally based on an *output* (rather than outcome) funding formulas, of which the simplest version is funding = *quantity x output unit cost*. Because accruals measures output costs better, it will immediately be obvious that FF will operate better with AA. However, this does not mean that all forms of cost-based FF necessarily require AA. Particularly if the funding formula funds variable costs (i.e. does not cover capital costs), some funding formulas may potentially work fine in a cash accounting and budgeting context. As a generalization, however, it may be said that cost-based funding formulas are likely to require accruals.

Key Readings

Robinson, M (2007), "Cost Information" in M. Robinson (ed.) *Performance Budgeting*.

Robinson, M (2009), *Accrual Budgeting and Fiscal Policy,* IMF Working Paper, obtainable at http://ideas.repec.org/p/imf/imfwpa/09-84.html.

Diamond, J (2002) *Performance Budgeting – Is Accrual Accounting Required?* IMF Working Paper, obtainable at http://www.imf.org/external/pubs/ft/wp/2002/wp02240.pdf.

17. Parliament and Performance Budgeting

The role of parliament in a government-wide performance budgeting system varies greatly depending on the political system, and in particular upon the role of the parliament vis-à-vis executive government in the formulation of the budget. This is therefore an area in respect to which we need to be very cautious in making generalizations. Broadly speaking, however, we can distinguish between two extremes. At one end of the spectrum are parliaments which have a major direct influence on the allocation of resources in the budget. At the other end of the spectrum are parliaments which, although they must legally approve the budget, have in reality little influence over its content, which is essentially determined by executive government.

The US Congress is a notable example of the first type of system. Although the American president presents a budget to Congress for consideration, there are few limits on the changes which Congress can (and does) make before voting its own final budget. In respect to the second type of system, the British parliament is the most striking example. In Britain, not only is the parliament usually tightly controlled by the cabinet via the party system, but any amendment by the parliament to the budget presented by the executive is considered to be a motion of no confidence, the passage of which would lead to the government's resignation. Parliament therefore essentially accepts the budget or sacks the government. The parliamentary systems of many countries lie in between these extremes, with the parliament exercising some degree of direct influence over the allocation of resources in the budget.

To the extent that the parliament has substantial independent budgetary power, a government-wide performance budgeting system will only work to insofar as the parliament links its funding decisions to results. If, however, the parliament has little or no independent budgetary power, this is essentially irrelevant. In either case, however, the *performance accountability* (as opposed to budgetary control) role of parliament is a potentially important part of a successful performance budgeting system. We therefore start by considering aspects of this accountability role.

Parliament and Performance Accountability

In successful government-wide performance budgeting system, *performance reporting* – that is, reporting to parliament on the objectives and results achieved by government agencies – is a key element of the system. A growing number of countries have adopted an approach originally adopted in the US under the 1993 Government Performance and Results Act, under which

parliament is presented with annual performance plans and annual performance reports for each ministry or sector within government.

Attempts to encourage parliament to focus on and debate performance in plenary session (i.e. when the parliament meets as a whole) have generally met with limited success (see box for example).

Disappointing Dutch Experience in Parliamentary Plenary Performance Review

In order to increase the political relevance of [the debate on annual performance reports], [the government] proposed in 2004 to hold a plenary debate on the annual reports on the third Wednesday of May in the presence of the entire cabinet (Blok Resolution 2004). During the May 2005 debate, the entire Cabinet was present. However, the floor leaders of different large parties were not present at the debate. The number of MPs present at the debate fluctuated between 35 and 60 (of 150) despite the fact that there were no parallel committee meetings . … The Blok Resolution …has led to an increased political attention for the annual reports and increased media coverage. The attention for the plenary debate unfortunately went hand in hand with a diminishing attention for the commission debates. The attention of members of parliament for the annual report remains rather low. Interest tends to be confined to the parliamentary finance commission. [50]

It is parliamentary committees which tend to be the main forum for effective parliamentary performance accountability. Approaches to enhancing the role of parliamentary committees in this area vary greatly internationally. One approach is to expand the role of the finance/budget committee to cover performance as well as financial issues. Building up the role of sector committees (i.e. committees covering specific sectors of government such as health) can, however, be very useful. As is true in other areas of parliamentary committee work, the degree of success of parliamentary committees in building their role in performance accountability tends to be closely linked to the support resources at their disposition – particularly in respect to professional research support staff. Particularly useful in this context is the creation of professional analytic services which support the parliament as a whole, of which the US Congressional Budget Office is one of the best-developed examples.

The supreme audit institution (e.g. national audit office, court of accounts) is in many political systems intended to serve as an accountability instrument of the parliament, independent of

[50] Source: Sterck & Bouckaert (2006), p. 14.

executive government. In this context, the extension of the role of the SAI from purely financial audit to performance audit (as discussed in the section of performance auditing) is potentially a useful means of reinforcing the performance accountability role of the parliament. With performance audits being presented to it by the SAI, the parliament is provided with a stronger information base with which to hold executive government to account for the results achieved with taxpayer resources.

Parliament and Expenditure Priorities

When parliament exercises substantial independent budgetary power, it becomes essential to the success of government-wide performance budgeting to encourage and facilitate it to use results information when considering the budget. Philip Joyce (2007) identifies certain important steps which can be taken to this end, including the presentation of performance information in a form meaningful and readily usable by parliamentarians and the establishment of a regular routine for program review and reauthorization. Beyond this, the scope for generalizations about how to promote performance budgeting in parliament is a matter which needs to be examined in a country-by-country context, with regard to key constitutional variable (e.g. does the constitution prevent the parliament from increasing aggregate expenditure – if so, this tends to force it to be more focused on prioritization) and political variables (e.g. how strong is the party discipline – to the extent it is weak, it becomes very hard to prevent budgeting fragmenting into un-prioritized expenditures favoring particular parliamentarians constituencies). Experience certainly suggests that, the greater the fragmentation of budgetary power in the political system – both between executive government and the legislature, and within the legislature – the more challenging it can become to implement an effective performance budgeting system.

Key Readings

Joyce, P. (2007), “Performance Budgeting under the separation of powers”, in Robinson (ed), *Performance Budgeting*.

Lienert, I. and M.-K. Jung (2004), *The Legal Framework of Budget Systems*, Paris, OECD, obtainable at http://www.oecd.org/dataoecd/3/28/43487903.pdf.

Mordacq, F. (2008), “Budgetary Reform and Parliament: The French Experience”, paper delivered at International Symposium on The Changing Role of Parliament in the Budget Process, Afyonkarahisar, Turkey, 8-9 October 2008, obtainable at http://www.sigmaweb.org/dataoecd/63/4/41831999.pdf.

18. Medium-Term Budgeting and Performance Budgeting

Performance budgeting has usually seen as being closely linked to medium-term budgeting. In this section we consider why this is the case. The starting point is a discussion of the nature and different versions of MT budgeting, the key technical challenges it involves, and the experience with MT budgeting in developing countries.

Nature and Benefits of MT Budgeting

The core elements of a MT budgeting are:

- A clearly stated aggregate MT fiscal framework (MTFF), including a statement of desired medium-term outcomes for key fiscal aggregates (especially deficits and debt)—that is, medium-term fiscal targets—and of the aggregate expenditure ceiling and aggregate revenue levels consistent with that fiscal policy.

- A "top-down/bottom-up" reconciliation process to ensure that expenditure and revenue policy are consistent with the MTFF—that is, to maximize the probability that the government's expenditure and revenue policies will, over the medium-term, result is levels of aggregate expenditure and revenue consistent with the fiscal targets set in the MTFF.

- Resulting from this process, a set of medium-term sectoral or ministry expenditure projections or ceilings which both embody government expenditure policies and are consistent with its MT fiscal policy. In most countries, these are indicative (i.e. the government makes no advance commitment to maintaining sectoral/ministry expenditure at the projected levels), but in some countries the government commits itself to funding sectors or ministries at these projected MT levels.

The key benefits of an effective MT budgeting are:

- Improved aggregate fiscal discipline: MT budgeting significantly reduces the chances of sound fiscal policy being undermined by higher-than-expected expenditure arising from expenditure policies and commitments which the government made in the past without fully considering their future implications.

- Reduced spending ministry uncertainty about their future levels of funding, leading to better planning and management. This reduced uncertainty is greatest when government makes commitments about funding over the MT time horizon. However, even where the MT sectoral/ministry expenditure projections are only indicative, effective MT budgeting can significantly reduce uncertainty. This is because, to the extent that the MT budgeting process ensures that all expenditure policy decisions are fully consistent with aggregate fiscal policy over the medium-term, it will substantially reduce the need for government to make *ad hoc* spending cuts because of unexpected fiscal developments.

- Improved expenditure prioritization: the process of reconciling expenditure policy with top-down aggregate expenditure ceilings encourages a more systematic consideration of expenditure priorities. For this reason, effective MT budgeting is closely linked with improved expenditure review and prioritization processes, of the type which are discussed in the section on expenditure prioritization. These processes provide an excellent means by which to feed performance information more systematically into budget preparation, providing an important link between the MT budget framework and performance budgeting.

Forward Estimates

A fundamental tool of MT budgeting is what is often referred to as "forward estimates". [51] These are MT estimates of expenditure and revenue on a "current policy" basis—that is, projections which indicate what expenditure and revenue will be in each of the next three or four years if there are no new spending initiatives, no changes to tax laws, and all explicit and clear commitments made to future expenditure (including political promises) are taken into account. Armed with good forward estimates, ensuring the consistency of "bottom up" expenditure and tax policy with "top down" aggregate expenditure ceilings driven by fiscal policy is a two stage process, involving:

- Firstly, an assessment via the forward estimates of whether unchanged expenditure policy is consistent with desired levels of aggregate expenditure. If the forward estimates indicate that existing expenditure policy will result in an excessive level of aggregate expenditure over the medium-term horizon, then adjustments are made to current spending policy (or,

[51] Also known by a range of other names, such as "annual reference level update" in Canada.

potentially, to tax policy). If, on the other hand, current expenditure policy implies expenditure below the aggregate envelopes, then there is room for new spending initiatives (or tax cuts).

- Secondly, a similar assessment of the impact of potential new spending initiatives, using estimates of their MT costs to determine whether they can be accommodated without breaching the aggregate expenditure envelope or whether they must either be scaled back or room found for them by cutting elsewhere.

Success in reconciling expenditure and revenue policy with sound fiscal policy over the MT depends critically on the quality of the forward estimates—that is, on the extent to which the forward estimates provide an accurate forecast of the levels of aggregate expenditure and revenue which will result over the medium term from current expenditure and revenue policies. Only to the extent that the forward estimates are accurate will they provide advance warning that changes in expenditure policy are needed to meet aggregate fiscal constraints.

Poor quality forward estimates will tend to undermine the entire MT budgeting process. Countries which have attempted to introduce MT budgeting without investing significant effort in the forward estimates process tend, unsurprisingly, to have been disappointed with the results. In the absence of a system and capacity to produce quality forward estimates, projections of medium-term aggregate spending and revenue tend to be prepared on the basis of the crudest techniques (e.g. updating based only on the application of a general inflation factor) which fail to capture the dynamics of current policy. This makes the top-down/bottom-up reconciliation process pointless, and it is therefore not surprising that such countries have also tended not to invest much effort in the enhancement of this process.

MT Ceilings: Fixed vs. Indicative?

MT budgeting produces a set of MT expenditure projections, often referred to as “ceilings”. As noted above, however, the status of these ceilings differs significantly between countries.

At one end of the spectrum are countries which set fixed MT ceilings for ministries. That is, the government makes commitments to ministries about the level of funding they can expect to receive over the coming, say, three years. One of the most well developed examples of this is the UK, which since the late 1990s has operated a system based on fixed three-year budgets for

spending ministries (recently extended to fixed commitments lasting the life of the five-year parliament). A number of other OECD countries – including France – have more recently followed the British example.

At the other end of the spectrum are countries where the MT expenditure projections are indicative and imply no commitment to ministries in respect to the MT funding they will receive. In such countries, government retains and exercises the right to change ministry and sectoral allocations each year, in order to progressively improve its expenditure prioritization. Australia is an example of this approach.

The Australian Approach to MT Budgeting

In the Australian system, the budget which each spending ministry receives is based essentially on the expenditure forward estimates — with modifications only for such new policy proposals as the government accepts (or, conversely, for cuts arising from explicit decisions to eliminate some programs inherited from the past). In preparing the budget for, say, 2011, the finance ministry first takes the estimates of 2011 expenditure which were prepared in 2009. It updates those estimates for any changes which have impacted on the costs of delivering existing policy, as well as any policy changes which might have happened during 2010, and then uses this as the starting point of the budget preparation process. Spending ministries are, approximately speaking, told not to bother debating the budget they need to deliver ongoing programs in the coming year – these are given by the updated forward estimates. Any requests for additional funding which the spending ministries make must, instead, be based on proposals for new policy initiatives. Thus the budget bids lodged by spending ministries early in the budget process – the so-called Portfolio Budget Submissions – are entirely focused on proposed expenditure policy changes. Using the expenditure forward estimates in this way creates a powerful incentive for the spending ministries to collaborate with the MoF in developing reliable forward estimates methodologies, because getting the estimates wrong could adversely affect the spending ministries budget position. In Australia, forward estimates were developed some years before they started to be used for budget preparation in the manner described above. It was only after they started, in the early 1980s, to be used as the basis for determining ministry budgets, that it became essential to substantially upgrade the quality of the estimates. This underlines the fact that this system can only work well once the quality of forward estimates has become sufficiently good that the estimates provide a broadly adequate assessment of the future costs of expenditure on an "unchanged policy" basis.

Both of these approaches represent legitimate models of MT budgeting. As noted above, even when ceilings are indicative only, spending ministry uncertainty about the levels of MT funding they will receive is reduced, because the compatibility of the indicative ceilings with aggregate fiscal policy will (unexpected shocks aside) have been verified via the MT budgeting process.

The technical requirements for setting fixed MT ceilings are considerably greater than for indicative ceilings. They include:

- Excellent macroeconomic forecasting,
- Excellent forward estimates of both revenue and expenditure,
- Reasonable macroeconomic stability,
- A strong expenditure review and prioritization mechanism,
- Sustainable aggregate fiscal policy settings.

Without a capacity to make good medium-term revenue forecasts, there is the danger that unrealistically high multi-year expenditure ceilings will be set. Down the track, when revenue turns out to be below forecast, the government will be faced with the difficult choice between withdrawing its budget commitments to spending ministries or finding additional revenue.

Without good processes to scrutinize and reprioritize expenditure, making medium-term budget commitments to spending ministries will simply increase expenditure inflexibility, and make it harder to reallocate spending over time to the areas where it is likely to deliver the greatest benefits. With only limited capacity to review expenditure, the MoF may find itself forced to recommend to the government multi-year ceilings largely based on the *status quo* rather than on critical analysis of the *status quo*. Under such circumstances, ministries which should have the budgets reduced could easily find themselves protected from scrutiny for a number of years at a time. It is highly relevant that in the UK's case, these expenditure ceilings are set only after thorough spending reviews (discussed in the section on expenditure prioritization).

MT Budgeting and Performance Budgeting

A MT budgeting perspective can greatly reinforce the effectiveness of a government-wide performance budgeting system. This is first and foremost because expenditure prioritization – the central objective of program budgeting systems – works best when carried out from a MT perspective. In other words, when deciding how best to allocate budget resources between programs, it is highly advantageous to know the cost of those programs not just for the next financial year but also for the following years. Program costing on an annual basis may be misleading. This is particularly true for new programs, for which the first-year costs are often significantly different from the ongoing costs (due, e.g., to substantial program set-up cost, or to the fact that the program may not be fully operational until part way through the financial

year). It is therefore essential that new program proposals be accompanied with MT cost estimates. The reduced uncertainty about future funding levels which MT budgeting provides helps support performance budgeting because decisions about program allocations can be made with greater confidence that they will be sustainable – in other words, that it won't be necessary to revisit them two or three years down the track because they can't be afforded.

Conversely, performance budgeting helps MT budgeting. Most fundamentally, program budgeting will enable the best quality expenditure forward estimates to be developed, because it will facilitate the modeling of program-specific cost drivers (that is, demand or cost factors which affect the evolution of the ongoing costs of specific programs). In a system of fixed MT ministry ceilings, performance budgeting is virtually a necessity because, as outlined earlier, excellent expenditure prioritization mechanism are critical. Thus in the UK system, good performance information (especially, but not only, excellent performance measures) has been an essential tool of the triennial spending reviews.

The Medium-Term Expenditure Framework (MTEF)

To this point, the discussion has been about MT budgeting in general, with no mention of the familiar concept of the medium-term expenditure framework (MTEF). The MTEF concept has been the cornerstone of the approach to MT budgeting advocated by international organizations including the World Bank and the International Monetary Fund. The problem, however, is that the term "MTEF" has been used in so many different ways by different authors that it can no longer be said to have a clear, widely-agreed definition. In particular, the term is very often used in a sense much wider than multi-annual budgeting – for example, to refer to a whole system of budget preparation, often incorporating also policy formulation and planning. [52] This extremely broad definition was promoted in the World Bank's 1998 *Public Expenditure Handbook*, in which the MTEF was defined as "*a whole−of−government strategic policy and expenditure framework* … The MTEF consists of a top−down resource envelope, a bottom−up estimation of the current and medium−term costs of existing policy and, ultimately, the matching of these costs with available resources". While the second part of this definition is essentially the same as the concept of MT budgeting presented above, the italicized words

[52] Indeed, the *Public Expenditure Handbook* could itself be seen as having contributed to this much broader interpretation. Notwithstanding the crisp definition of an MTEF quoted above, there are a number of comments elsewhere in the manual which suggest a much more expansive notion of the MTEF.

could be read to imply a definition which covers not only the whole of the budget preparation process, but the planning process as well. On the other hand, one can find some authors who use the term MTEF in a much narrower sense – for example, refer only to detailed forward expenditure estimates. [53] The term "MTEF" is not one which is used much by OECD countries, where the terms "multi-annual" or "medium term" budgeting are much more common.

The broad conception of an MTEF as a total budgeting, planning and policy system means that the MTEF concept has often been seen as including within it, amongst other things, performance budgeting and the expenditure prioritization mechanisms. The problem with this is that it is really not possible or desirable to specify as part of the definition of an MTEF specific forms of performance budgeting and specific types of expenditure prioritization mechanism which are appropriate to all countries and which will be agreed for all time to be the best approaches. Thus, for example, many (including the *Handbook*) say that a sector-wide approach to planning is an essential element of the expenditure prioritization mechanism under an MTEF. However, as discussed in the expenditure prioritization session, there are good grounds to dispute the effectiveness of this approach to prioritization.

The same point applies to performance budgeting. Some proponents of MTEFs have seen it as necessarily incorporating program budgeting. Others have taken the view that some form of "output budgeting" or "activity-based budgeting" – by which they mean a formula funding approach to estimating budget requirements as a function of outputs or activities to be delivered – is a necessary part of an MTEF. This type of "output budgeting" (which is critiqued in the section on program costing and accounting) was fashionable in the late 1990s and early 2000s, but is much less so today. Again, it is not clear that it makes sense to incorporate such specifics into the definition of an MTEF.

While discussing the MTEF concept, two other widely used acronyms – Medium-Term Fiscal Framework (MTFF) and Medium-Term Budgeting Framework (MTBF) – should be mentioned. The term MTFF, to which reference was made at the outset of this section, has a clear meaning. It is widely understood to refer to the aggregate fiscal policy element of an MTEF – clear medium-term fiscal objectives, and a set of fiscal projections consistent with them. The term

[53] This appears to be the version of the concept used by Allen & Tommasi (2001), p. 180. For Allen and Tommasi, an MTEF so defined is merely a component of a broader "medium-term budget framework".

MTBF has, by contrast, been used in so many radically different ways that it is pointless to attempt define it. [54]

Overall, it seems to make little sense to approach MT budgeting based on definitions – in other words, to start with a definition of what an MTEF is supposed to be, and then work on introducing an MTEF based on that definition. It is much more important to focus on the core elements of MT budgeting, and the optional extras (e.g. fixed MT ministry ceilings) and work out on a country-by-country basis what elements will work best.

At the time of writing, considerable effort was underway in international organizations to re-think the whole issue of MT budgeting, and the MTEF in particular, and it appeared likely that a new paradigm would emerge within the subsequent years.

Developing Countries and the MTEF

The MTEFs introduced in developing countries have differed greatly, particularly in terms of the "add-on" elements of prioritization mechanisms and performance budgeting. The core element has been the development of MT ministry expenditure "ceilings". While there have been different approaches on the question of whether these should be regarded as fixed or indicative – and whether they would be rolling or fixed – it has been universally assumed that the MT ceilings should at the very least guide annual budget preparation.

In practice, the experience with MTEF has often been quite disappointing. The biggest problem has been that the funding allocations in the annual budget have often had little or no relation to the MT ceilings. Expressed differently, the MTEF has been a document with little impact on the real budget. This has particularly tended to be the case in countries where the MTEF was essentially prepared by consultants, with limited ownership by the authorities, so that the expenditure priorities expressed the MTEF were not necessarily priorities to which the authorities are really committed.

Poor quality forward estimates have been perhaps the most widespread obstacle to the development of effective MT budgeting in developing countries. As a South African commentator (Fölscher, 2007) noted in respect to the experience with MT budgeting in African countries:

[54] See, for example, the different concepts of an MTBF in the IMF's *Fiscal Transparency Manual* (2007: 129) and in Allen and Tommasi (2001), p. 180.

> *…the quality of forward estimates is poor. They consist far too frequently of the proposed budget for the first year of a multi-year framework, followed by inflation adjusted projections of cost for the outer year ...they pay little attention to, for example, the likely phasing of policy implementation, changes in demand that will effect spending unevenly or the impact of once-off capital spending on the base-year estimates. ...A key aspect of embedding a medium-term perspective therefore is deciding what the rules are for rolling over and adjusting and determining the forward estimates.*

A key reason for this is that the preparation of forward estimates is, at the technical level, a more difficult task than is often assumed.

With all the focus on MT budgeting, it seems sometimes to be forgotten that many developing countries have major difficulties with the quality of their annual budgeting. Many are not able to prepare good quality *annual* projections of the funding requirements of ongoing services. Some question, in this context, whether it makes sense to be pushing countries to introduce MTEFs when the quality of their annual budgeting needs such considerable improvement. In this context, the findings of a recent review by the World Bank's Independent Evaluation Group (IEG) are highly relevant. The IEG (2008: xv) noted the importance of "dealing with the basics first, such as ensuring that the government is executing a one-year budget reasonably well before launching sophisticated multiyear budgeting."

It may therefore make more sense for developing countries to think of the task not so much as one of "introducing an MTEF", but rather as one of gradually introducing more of a MT perspective into the budget, starting initially with efforts to define clearer medium-term fiscal objectives and gradually improve the quality of expenditure and revenue forecasts.

If it is the case that the introduction of full-blown MT budgeting is a difficult task which not all developing countries can aspire to in the short term, what does this mean for performance budgeting in developing countries? It would be difficult to argue that a full-blown MT budgeting system is an essential prerequisite for performance budgeting. Even in the context of a strictly annual budgeting system, introducing the systematic use of performance information can help to significantly improve expenditure prioritization and to put more pressure on ministries to improve the effectiveness and efficiency of their spending. From this perspective, one can think of performance budgeting and medium-term budgeting as two dimensions of PFM reform in

developing countries which have important synergies, but which should not be thought of as absolutely essential to one another.

In conclusion, introducing a MT perspective into budgeting can offer great benefits in terms of improved fiscal policy implementation and reduced funding uncertainty – and consequently better management – at the ministry level. MT budgeting has considerable synergies with performance budgeting: each reinforces the other in major ways. There is no single "correct" model of MT budgeting, nor even a single model of the MTEF. The form, and degree, of MT budgeting should be tailored to the circumstances of each country.

Key Readings

World Bank (1998) *Public Expenditure Handbook*, obtainable at http://www1.worldbank.org/publicsector/pe/handbook/pem98.pdf.

Kasek, L. and D. Webber (2009), *Performance-Based Budgeting and Medium-Term Expenditure Frameworks in Emerging Europe*, Warsaw: World Bank, obtainable at http://siteresources.worldbank.org/INTECA/Resources/WBperformanceBudgetingTEF.pdf.

Le Hourerou, P. and R. Taliercio (2002), *Medium Term Expenditure Frameworks: From Concept to Practice: Some Preliminary Lessons from Africa*, World Bank, Africa Region Working Paper Series, No. 28, obtainable at http://unpan1.un.org/intradoc/groups/public/documents/UNPAN/UNPAN002860.pdf.

Wynne, A. (2005), *Public Financial Management Reforms in Developing Countries: Lessons of Experience from Ghana, Tanzania and Uganda*, African Capacity Building Foundation Working Paper no. 7, obtainable at http://unpan1.un.org/intradoc/groups/public/documents/UNPAN/UNPAN002860.pdf.

19. Implementing Government-Wide Performance Budgeting

This section summarizes key implementation issues which arise in implementing a government-wide system of performance-informed budgeting. In other words, the focus is on the implementation requirements of a program budgeting system which aims to make systematic use of performance information in budget decisions to improve expenditure prioritization and increase pressure for performance. The special implementation requirements of more complex systems of performance, such as purchaser-provider, are discussed elsewhere in this manual.

The development of a *performance information base* is, obviously, one of the first steps in the development of a basic system of performance-informed budgeting. The development of program performance indicators should not be thought of as a single-step reform which can be accomplished in a short time-frame such as a couple of years. Rather, it is a gradual process which – as experience in leading OECD countries shows – will proceed over decades. This does not, however, mean that it is necessary to wait decades before a workable performance budgeting system can be introduced. Even a highly selective set of program indicators can enormously improve the expenditure prioritization decisions made by the government with the support of the finance ministry. The initial aim should therefore be to develop a relatively small set of useful program indicators, along with relatively simple program evaluation which is designed to be usable in the budget process.

It is very important not to neglect evaluation in a single-minded effort to develop indicators. Indicators in isolation are sometimes of limited value, and evaluation can be crucial to interpreting indicators. In the early stages of the developing of a government-wide performance budgeting system, the focus will tend to be on the development of simple – in many cases “desk based” – reviews of a small set of key programs during each budget cycle.

The development of a program structure for the budget is a reform which takes some years. The first step is the definition of appropriate programs in spending ministries – which means programs which are defined in terms of outcomes and outputs (with limited exceptions such as administration programs). This can usually be expected to take an absolute minimum of two years (and usually longer than that). In this process, spending ministries need extensive support from the MoF, in the form of technical instructions and guides and a major training program. The actual development of the programs needs, moreover, to be an interactive process in which the spending ministries prepare drafts and these are then reviewed and approved by the

finance ministry. It is, on the one hand, undesirable that the MoF simply impose its choice of programs upon spending ministries – not only does it not have sufficient knowledge to do this properly, but the lack of "ownership" of the resultant program structure by the spending ministry will tend to undermine the reform. On the other hand, because the programs are intended to be an instrument of central expenditure prioritization, spending ministries cannot be left to just define programs in any way they wish.

The modification of the accounting system and the financial management information system (FMIS) to make them program-friendly is a quite major implementation step. Although changing the chart of accounts is the first step, modifying the FMIS is usually the biggest task here. It may not make sense to wait till the introduction of a completely new IFMIS to introduce a program structure to the budget, so modifications to existing systems are often what is needed. Extensive training in new accounting procedures and related systems changes will be required.

Once a program structure has been developed and the accounting system and FMIS have been made program friendly, it is possible to start approving the budget on a programmatic basis. In the run-up to this, most countries first develop indicative program budgets which are presented to the parliament as an annex to the traditional budget. Such an indicative program budget shows the parliament and public what the budget appropriations would look like if they were approved in programmatic terms.

In moving to program appropriations, it is crucial to make a clear decision about the manner and extent to which traditional budget controls will be reduced. Most important here is the question of the degree of line-item control simplification. It is a serious mistake to simply impose program appropriations on top of a highly detailed traditional budget – this will not only deprive agencies and managers of the increased managerial freedom which is an essential part of performance budgeting, but will in fact tie their hands even more than was previously the case. In reducing line-item controls, most countries will not wish initially to go as far as OECD countries like France, Australia and the UK. But they will usually need to significantly reduce the line-item controls which they have previously enforced.

As discussed in the sections on PFM reform and expenditure prioritization, an important part of the implementation of a successful government-wide performance budgeting system is the reform of the budget process and the reduction of expenditure rigidities so as to make it possible to more readily reallocate limited public resources to the sectors and programs where it can deliver greatest social benefit.

The MoF will necessarily play a pivotal role in the implementation and ongoing operation of the performance budgeting system. Changes in the *modus operandi* and skill set of the MoF are therefore essential if performance budgeting is to be a success. The MoF can no longer be simply a controller of expenditure. It must play a central role as an expenditure policy adviser to the government. In particular, it must develop its capacity to advise government on where expenditure can be cut – e.g. via cuts to unsuccessful or low priority programs or efficiency improvements – in order to make room for new priorities or to assist fiscal consolidation. Such a role will in most countries requires a broadening of the skill base of the ministry, with the reinforcement of policy analysis skills alongside accounting and economic analysis capacity. This will normally require a combination of retraining and new recruiting. The increased expenditure policy role of the MoF will require additional resources, which raises the question of overall staffing levels. However, part of the answer to this can be found in the systematic elimination of unnecessary traditional control functions, such as the review and approval of large numbers of transfer requests arising from excessively detailed line-item controls.

In discussing here the role of the MoF, it should be recognized that in some countries the expenditure policy role described above will appropriately be shared with other central agencies (e.g. the prime minister's or president's office).

Broadly speaking, countries may choose between a "big bang" and a more gradual approach to the implementation of performance budgeting. The "big bang" approach – of which Russia provides one example – means implementation within a couple of years. Generally speaking, it will require that the country concerned already has not only the major PFM prerequisite in place, but that the FMIS works well and can be relatively easily adapted to programs. The great advantage of the big bang approach is momentum. By contrast, a gradual approach may be more realistic in many countries. Equally, however, the great danger of too gradual an approach is the loss of momentum, as a result of which the reform may never be completed.

The example of France is a good one to have in mind when determining the implementation timetable. After having spent some years determining the broad parameters of its new performance budgeting system, France deliberately chose a five year implementation period between the 2001 passage of the law mandating the new system and its coming into full force in 2006. This five year period proved to be an extremely busy one, give the scale of change which the new system involved, and one could not say in retrospect that the reform could have easily been implemented in a shorter period of time.

The management of the implementation process is critical to success. Performance budgeting cannot, firstly, be implemented successfully without strong support from the political

leadership. It will usually make sense for the MoF – perhaps in conjunction with other central agencies – to create an implementation task force. Spending ministries need to be closely involved in the process at a senior level, perhaps via a consultative committee comprised of senior ministry representatives. Throughout, the MoF will need to provide strong technical guidance to spending ministries to assist them in properly developing the new systems (e.g. properly defining programs and selecting the right types of program performance indicators).

Developing countries often face more serious implementation challenges arising, for example, from greater capacity constraints, weaknesses in their overall PFM systems and, sometimes, from governance problems which increase expenditure rigidities. These have to be taken into account when determining the target timeframe for implementation. The resource and capacity constraints which face developing country MoFs make it particularly important to simplify unnecessary traditional controls, in order to free staff time for new responsibilities. Above all, however, developing countries can ease the implementation burden by keeping their performance budgeting systems simple. In this context of program budgeting, this means, for example, avoiding overly complex program structures and unnecessary "add-ons" such as activity-based costing. More generally, it usually means deferring to the future the possible application the more sophisticated performance budgeting mechanisms such as formula funding and purchaser-provider. Performance targeting should also be approached in a gradual and highly selective manner. Overall, the initial focus should be simply on making the budget more *performance informed*: that is, on introducing a systematic consideration of program results when determining the allocation of resources in the budget.

In conclusion, performance budgeting should not be seen as a simple reform to be introduced overnight. Even in its least complex form, it is a demanding multi-faceted reform which takes years to put in place.

Key Readings

Diamond, J. (2007), "Challenges to Implementation", in Robinson (ed.), *Performance Budgeting.*

20. The Managing for Results Context

Performance budgeting uses the budget as a tool for making public management more results-focused. It is, however, only one of a range of "managing-for-results" instruments which share this broad objective. In this final section we look at some key areas of MFR reform which need to be pursued together with performance budgeting if the full benefits of the latter are to be realized.

Civil service reform is one of the most crucial accompanying reforms in many countries. There are a range of reasons why, to varying degrees in various countries, the civil service may not be sufficiently client and results-focused. MFR seeks to change this by focusing employment arrangements – hiring, pay, promotion and termination – more on results. Formal performance planning and monitoring is a key part of this, and makes increasing use of tools such as formal performance "contracts" or agreements.

It is in respect to the setting of clear objectives for organizational units and individuals within them, and the development of performance measures to assess the degree of achievement of these objectives, that civil service performance management needs to be linked to the broader strategic planning and performance budgeting systems. In particular, the objectives and indicators established for work unit and individuals should be linked to program and ministry objectives and indicators. Generally speaking, objectives and indicators should "cascade" through organizations, linking those for the organization as a whole to those set for individuals and work units.

The question of performance incentives is an important one for performance budgeting. In traditional civil services, the material incentives for good performance are often weak, and it follows that strengthening these – for example, through performance pay arrangements – can be an important element of MFR. However, we need to be aware of the potentially perverse effects which too large a component of performance pay may create as a results of the inevitably imperfect measurement of individual or work unit performance. It is also a major mistake to view performance motivation as exclusively a question of material incentives. "Public service motivation"—the altruistic commitment of public workers to client service (e.g. doctors and nurses to patients) or to a cause (e.g. the desire of environment ministry staff to save the environment) – is very important. An important part of good performance management is the reinforcement and harnessing of public service motivation.

Managerial flexibility in human resource management is a very important part of MFR. Traditional civil service arrangements tend to give managers little scope to reward good staff, or to sanction or terminate ineffective staff. In countries which have gone furthest along the MFR path, managers have acquired much greater responsibility and authority in these areas. However, the scope for movement in this direction has to be viewed in a country-specific context, taking into account the overall status of governance – for the same reasons as were discussed in respect to employment flexibility in the section on PFM reform.

Finally, the success of MFR depends completely on the political circumstances of the country and, more specifically, on the degree to which the politicians and the electorate are themselves performance-focused. If, for example, the political system is one where continuing electoral support for government has little relationship to their capacity to deliver efficient and effective expenditure – but is more related to, say, group loyalties or rent-seeking behavior – performance budgeting, and MFR more generally, are not likely to make much headway.

21. Concluding Comments

Performance budgeting is a challenging but potentially important reform in the ongoing struggle to make government more results-oriented. Coupled with other "managing for results" reforms, it can help substantially improve the effectiveness and efficiency of public expenditure. The objective of this manual has been to provide an outline of key forms of performance budgeting, their differing performance information requirements, the key design and implementation issues which they raise, and their pre-requisites and co-requisites. The diversity of performance budgeting models has been emphasized. It is important for each country to make a clear decision about what type of performance budgeting which is appropriate to its own circumstances. This decision should explicitly take into account the specific capacity and other characteristics of the country concerned. Another key point emphasized has been the importance of a realistic approach to implementation strategy and timetable. Successful performance budgeting is not something to be introduced in a rush. It is, rather, a reform which can be expected to take at very minimum several years to implement in a preliminary way, to reach the point where budgeting starts to become more systematically "performance-informed".

Glossary

Accrual accounting	System of accounting in which a key role is played by the recording of *expenses*, assets and liabilities.
Accrual budgeting	Budgeting system in which budget approvals take the form principally of authorizations to ministries to incur *expenses.*
Accrual output budgeting	Purchaser –provider budgeting model applied to the whole government budget, as introduced in New Zealand and Australia in the 1990s.
Activities	Types or categories of work undertaken in the production and delivery of outputs
Activity-based costing	A costing methodology in which input costs are allocated to categories of activity, using an allocation basis which reflects, as closely as possible, the real consumption of resources by those activity categories. When used for program costing, the activity costs are allocated between programs in a second stage.
Aggregate fiscal policy	The government's overarching objectives for the budget deficit, debt and other relevant fiscal aggregates.
Appropriation	Amount of expenditure legally authorized by the parliament to be undertaken.
Administration program	Programs which cover overhead costs of a ministry or agency, such as central management and personnel services. Also known as, e.g., administration programs or support programs.
Allocation basis	Formula or principle used to allocate specific expenditure item between two or more "cost objects"—in the context of program costing, between two or more programs.
Alternative budgeting	A variation of zero-base budgeting in which decisions are focused not on a zero base but on the margins near the current budget base. Usually three or more alternative budgets have to be submitted for each program. Generally, at least one of the alternatives has to be less than the current budget. Often a specific percentage reduction is mandated.
Bottom-up budgeting	Budget preparation process which is based primarily on spending ministries making bids for resources, with central consideration of those bids then determining the overall shape of the budget.

Budget classification	Categories of expenditure used in the budget, particularly for the approval of expenditure.
Budget preparation	Stage of the budget process the government decides how much funding it will provide to which agencies and for which purposes, and prepares the budget law which will be presented to parliament for its approval.
Budget execution	Implementation of the expenditure plan developed in the budget – including the entering of contracts and expenditure of funds.
Cash accounting	A system of financial accounting in which the focus is upon money actually paid out (expenditures or outlays) and money actually received in the financial year.
Chart of Accounts	System of classification of transactions – revenue, expenditure, financing etc – used for accounting purposes.
Classification of the Functions of Government (COFOG)	Functional classification of expenditure developed by the United Nations and incorporated into the International Monetary Fund's Government Financial Statistics methodology.
Commitment	Contractual obligation to make a future payment.
Control total	A centrally-imposed quantitative limit on some category of expenditure which can be undertaken by the spending ministry during the budget year. May be imposed by the parliament or by a central executive body such as the cabinet of ministers or the finance ministry.
Cost-effectiveness	The achievement of intended outcomes at lowest possible cost.
Cost driver	In the context of program budgeting, a formula or principle which is used to determine the proportions in which an indirect cost is split between two or more programs (or sub-programs) to which it contributes.
Direct cost	In the context of costing programs, expenditure on inputs which contributes to only one program.
Effectiveness	The degree of success of an output in delivering its intended outcomes.
Efficiency	Production of an output at minimum cost while holding quality constant, given prevailing input prices.
Evaluation	Analytic assessments typically addressing the cost-effectiveness or appropriateness of public policies, organizations or programs.
Expense	Costs attributable to the present financial year in the accrual accounting framework. Represents a consumption of resources,

	regardless of whether it is paid for immediately or results instead in a liability or a reduction in assets.
External factors	Factors outside the control of government which influence the outcomes achieved by public programs. External factors may be either client/case characteristics or aspects of the context in which the program is delivered. Sometimes also referred to as "contextual factors."
Financial management information system (FMIS)	Computerized systems for managing budgeting and associated processes. See also *IFMIS.*
Formula funding	When used as a performance budgeting tool, formula funding is a system in which funding provided by government to a public sector agency is an explicit (that is, algebraic) function of measures of expected and/or actual results—that is, of measures of outputs and/or outcomes.
Forward estimates	Projections by a central budget agency of aggregate expenditures over a fixed term (usually budget plus three years) on a no policy change basis. Forward estimates are the basis for medium term fiscal planning and may also be used as the basis for imposing expenditure limits.
Functional classification	Classification of expenditure by "functions", a somewhat imprecise term which represent a mix of output types and activity types.
Heterogeneity	The deliberate variation in the amount and/or types of activities delivered to different clients/cases receiving the "same" service, particularly in response to difference in client/case characteristics. For example, more intensive teaching activity directed to students with disabilities.
High-level outcomes	The more indirect outcomes of outputs, which arise as a consequence of the achievement of proximate outcomes. For example, in education, the proximate outcome of higher numeracy and literacy contributes to the higher-level outcome of better economic performance. Also sometimes known as "end" or "ultimate" outcomes.
Integrated financial management information system (IFMIS)	A computerized system which to a greater or lesser degree integrates multiple functions into the same system, so that accounting, expenditure control, payments, budget preparation and a range of other functions will be built into the same big computerized system.
Impacts	Term used by some to refer to longer-term or higher-level outcomes.
Incrementalism	Budgeting is characterized by "inattentiveness to the (budgetary) base"—in other words, that budgetary decision-makers take the budgetary base more or less for granted as the starting point in budget

	formulation, and focus their attention primarily on the size of the increment (or, occasionally, decrement) in agency or program budgets, mainly by a process of adjusting budgets for cost changes.
Indirect cost	In the context of costing programs, costs of inputs or activities which contribute to more than one program. More generally, shared costs which need to be allocated between a number of "cost objects."
Inputs	Resources used in the carrying out of activities to produce outputs (for example. labor, equipment, buildings).
Intermediate outcome	The more direct or immediate impacts of outputs. For example, in education, the student knowledge (for example, higher numeracy and literacy) is a key proximate outcome.
Line-item budgeting	Budgeting in which agencies are provided with budget appropriations specified in terms of input categories (that is, by economic classification).
Managing-for-results	The use of formal performance information to improve public sector performance across the board, including in human resources management, in strategic planning and budgeting.
Medium-Term Expenditure Framework (MTEF)	Not a concept with a clear generally-accepted definition. However, defined in World Bank *Public Expenditure Handbook* (1998) as "a whole−of−government strategic policy and expenditure framework … The MTEF consists of a top−down resource envelope, a bottom−up estimation of the current and medium−term costs of existing policy and, ultimately, the matching of these costs with available resources".
Medium-term Fiscal Framework (MTFF)	Explicit policy on desired medium-term outcomes for key fiscal aggregates (especially deficits and debt)—that is, medium-term fiscal targets—possibly supported by approximate medium-term forecasts of aggregate expenditure and revenue.
Outcome	Changes brought about by public interventions upon individuals, social structures, or the physical environment.
Outcome variable	The characteristics of individuals, society or the physical environment which public programs seek to change.
Output	A good or service provided by an agency to or for an external party.
Performance audit	Evaluation of the efficiency and effectiveness of public expenditure carried out by a supreme audit authority or other group of auditors.
Performance indicator	See *performance measure.*

Performance measure	Ratings or quantitative measures which provide information on the effectiveness and efficiency of public programs.
Performance target	Level of measured performance to be achieved by a specific date.
Perverse effects	Behavioral distortions or other adverse consequences of performance measures.
Processes	The processes by which inputs are transformed into outputs. Same as *activities.*
Program Assessment Rating Tool (PART)	Technique developed by the US under the Bush presidency in order to rate program performance for Tool use in the budget process. Programs rated in five categories ranging from "effective" down to "ineffective" and "results not demonstrated."
Program budgeting	The systematic use of performance information to inform decisions about budgetary priorities between competing programs, based on the program classification of expenditure (see programs).
Program hierarchy	Multi-level program-based classification of expenditure, in which programs are broken into one or more lower-level categories (sub-programs etc).
Programs	Categories of expenditure based on groups of outputs (or support services) with a common objective, which is usually an outcome.
Public Service Agreements (PSA)	System developed in the United Kingdom from the late 1990s until 2004 under which high-level performance targets are set for each ministry as part of a biennial spending review process which defines multi-year agency budgets. Targets have evolved over time from being primarily output focused to primarily outcome focused.
Purchaser-provider	Funding systems under which agencies are paid "prices" for systems the results (usually outputs) which they deliver.
Quality	The extent to which the characteristics of an output—in the case of a service output, the activities delivered and their timeliness— are such as to increase its potential capacity to achieve its intended outcome. Not to be confused with the outcome itself.
Spending review	The systematic scrutiny of existing expenditure to identify, in particular, options for cuts. It involves both program reviews (the review of specific services provided by government) and efficiency reviews (which focus on reducing the cost of delivering services).
Standardized outputs	Outputs where every client receives essentially the same level of service – that is, outputs which are not characterized by *heterogeneity*.

Strategic phase (of the budget preparation process)	Stage early in a well-designed budget preparation process where the political leadership sets the priorities which will guide the process.
Supreme audit institution	Generic term used internationally for bodies such as the UK National Audit Office, the US Government Accountability Office, and the French *Cour des Comptes*, which are independent (to varying degrees) of executive government and designed to hold it to account.
Transfers	Shifting of funds between appropriation categories.
Top-down budgeting	The unilateral setting of expenditure ceilings by the center (president, cabinet, finance ministry etc) without "bids" by the spending ministry.
Virement	See *transfers.*
Zero-base budgeting	A system of performance budgeting in which expenditure is broken down into, and analyzed in terms of, "decision packages" (also known as "service increments") which constitute a series of optional funding levels from the presumed base of zero to and beyond the current level of service. Priority rankings are attached to these decision packages, and these rankings are used to ensure that the available level of revenue funded those decision packages which are of highest priority.

References [55]

Allen, R. & D.Tommasi (2001), *Managing Public Expenditure: a Reference Book for Transition Countries*, Paris: OECD.

Bevan, G and C. Hood (2006), "What's Measured is What Matters: Targets and Gaming in the English Public Health Care System*, Public Administration*, 84(3).

Danish Ministry of Finance, 2006, *Introduction of a Cost-Based Appropriation System*, Copenhagen: Ministry of Finance.

DFID (2001), *Understanding and Reforming Public Expenditure Management,* version 1, obtainable at http://webarchive.nationalarchives.gov.uk/+/http://www.dfid.gov.uk/documents/publications/pfma-pem.pdf.

Fölscher, A. (2007), *Are We Asking the Right Questions? Embedding a Medium-Term Perspective in Budgeting*. Proceedings of a CABRI seminar on medium-term budgeting held in Ghana December 2007.

France (2004) *Démarche de Performance* : S*tratégie, objectifs, indicateurs. Guide méthodologique pour l'application de la loi organique relative aux lois de finances du 1er août 2001*, Paris: Le Ministre d'État, Ministre de l'économie, des finances et de l'industrie, et al.

HM Treasury (2002), *Resource Accounting and Budgeting Green Book,* London: HM Treasury

Hood, C. (2006), "Gaming in Targetworld: the Targets Approach to Managing British Public Services", *Public Administration Review,* July/August.

Independent Evaluation Group (2008), *Public Sector Reform: What Works and Why*? Washington: World Bank.

Kelman, S. and J. Friedman (2007), *Performance Improvement and Performance Dysfunction: An Empirical Examination Of Impacts Of The Emergency Room Wait-time Target In The English National Health Service*, Working Paper, Kennedy School of Government, Harvard University.

Potter, B. & J. Diamond (1999) *Guidelines for Public Expenditure Management*, Washington: IMF.

Robinson, M. and J. Brumby (2005) *Does Performance Budgeting Work?* Washington: IMF.

[55] Note: this list is confined to materials cited in the manual, and does not cover the key readings identified for each chapter.

Schiavo-Campo, S. & D. Tommasi (1999), *Managing government expenditure*. Manila: Asian Development Bank.

Sterck, M. & G. Bouckaert (2006), "The impact of performance budgeting on the role of parliament: a four-country study", Paper presented at the 2nd Transatlantic Dialogue, Leuven, June 1-3.