Responsibility budgeting and accounting

L.R. Jones,⁎ Fred Thompson

⁎Department of Systems Management SMJN, Naval Postgraduate School, Monterey CA 93943 USA
⁎Atkinson Graduate School of Management, Willamette University, Salem OR 97301 USA

Received 3 March 2000; accepted 23 November 2000

Abstract

Information/transaction costs make it necessary to decentralize some decision rights in organizations and in the economy. Decentralization in turn requires organizations to solve the control problem that results when self-interested persons do not behave as perfect agents. Capitalist economies solve these control problems through the institution of alienable decision rights. But because organizations suppress the alienability of decision rights, they must devise substitute mechanisms that perform those functions. Three functions are critical: (1) allocating decision rights among agents in the organization, (2) measuring and evaluating performance, and (3) rewarding and punishing individuals for their performance. Responsibility budgeting and accounting systems are the most widespread mechanisms for performing these functions in business today. © 2001 Elsevier Science Inc. All rights reserved.

Introduction

In this article we explain the nature of responsibility budgeting, its intellectual justification, its antecedents, and its present and future use in the public sector. This is not a straightforward task. We cannot simply explain how responsibility budgeting is used and how it works. Responsibility budgeting makes sense only as a part of a framework of structural, procedural, and monitoring/reporting relationships. We must, therefore, also explain the framework that gives it utility and power. At the same time, responsibility budgeting and accounting, or their functional equivalents, make an essential contribution to the efficacy of this broader framework of relationships. One cannot arbitrarily mix and match administrative relationships and expect that the outcome will be productive. The efficacy of
administrative relationships depends upon their congruity with each other as well as with the purposes and products of the entity in question and the productive and information processing technologies available to it.

**Governance arrangements, administrative processes, contractual relationships**

All governance arrangements and administrative processes are primarily mechanisms for motivating and inspiring people, especially subordinate managers, to serve the policies and purposes of the organizations to which they belong. This means that all governance arrangements and all administrative processes can be treated as contractual relationships and that administrative design and implementation can be thought of as negotiating and enforcing contracts.

One way of describing contractual relationships involves the language of principal and agent. This language implies a hierarchical relationship, in which a nominal subordinate (agent) serves the purposes of a superior (principal). On the presumption that behavior is largely self-interested, principal-agent relationships are problematic (give rise to agency costs) only where (a) the efforts of the agent cannot be perfectly observed; (b) the interests of agent and principal diverge; and (c) agents pursue their own interests, i.e., behave opportunistically.

One of the key goals of governance arrangements and administrative processes is the minimization of agency costs. Of course, agency costs also include all resources used to reduce divergences of interest, i.e., identifying collectively beneficial relationships, negotiating contributions, and devising procedures for monitoring performance and sanctioning defectors. Included here are a whole panoply of activities extending from the employment of security guards to the design and implementation of new or reconfigured accounting and reporting systems. Hence, minimizing agency costs means minimizing the sum of costs that result from opportunistic behavior plus the costs of avoiding or controlling that behavior (Zimmerman, 1995). Economic theory tells us that we find this optimum where the marginal costs of controls equal their marginal benefits, as shown in Fig. 1 (from Breton and Wintrobe, 1975).

Traditional or Weberian bureaucracies rely on rules to govern or prevent opportunistic behavior. In other words, principals specify in detail what agents must do (or must not do), carefully monitor their actions, and sanction all deviations accordingly. The problem with this approach is that agents often have better information about some things than do principals. Principals hire agents because of their superior expertise and to spare themselves the burden of being perfectly informed about every aspect of an organization’s operations. In neither case will principals have the knowledge needed to specify in detail what the agent should do without thereby sacrificing performance. This means that rules are not always a wholly satisfactory solution to the principal-agent problem. It is this fact that makes the application of agency theory to the public sector especially important, for it is in the public sector that the opportunity costs arising from detailed rules often seem highest.

Organizational economists do not generally advocate more rules as a way to control opportunistic behavior. Rather, they stress two alternative approaches. One is to improve
principals’ abilities to monitor agents. This is often referred to as improving “transparency.” For example, full accrual accounting gives a truer picture of resource use than does standard government accounting and thus helps make government operations more transparent. The second is to seek ways to align the incentives of agents with principals’ interests. This is the preferred approach of organizational economists and managerial accountants. As the New Zealand Treasury (1987) observed, “Incentives matter . . . . Well designed policies will align the interests and actions of individuals with those of the nation.”

A historical digression: from traditional bureaucracy to the M-form organization

Most large-scale organizations in the American public sector are organized like turn-of-the-century railroads. Operating responsibility is delegated on a geographic or site basis, rather than a line of business basis. Regional chiefs report to an agency head. Small armies of administrative staff specialists also report to agency heads. Their job is to gather and process quantities of data for agency heads to use to coordinate activities, allocate resources, and set strategy.

These structures can be traced directly to the administrative system developed by the Prussian bureaucracy under Heinrich von Stein, Gerhard von Scharnhorst, August von Gneisenau, and Helmuth von Moltke. The Prussian system included administrative innovations such as detailed centralized resource requirements planning (discretionary expense budgets), control by rules and standard operating procedures, functional organizational design, vertical integration, decomposition of tasks to their simplest components, sequential processing, and administrative centralization and specialization of administrative staff functions such as reporting, accounting, personnel, and purchasing.

The Prussian administrative system was once widely emulated by forward-looking businesses and governments all over the world. In the US, among the first large-scale organizations to adopt this system were the railroads and the military departments. In industry,
early adopters of the elements of the Prussian administrative system consistently grew large as hierarchy and bureaucracy created massive economies of scale and scope. Economies of scale are produced by spreading fixed expenses over higher volumes of output, thereby reducing unit costs. Economies of scope are produced by exploiting the division of labor—sequentially combining highly specialized functional units in multifarious ways to produce a variety of products (Chandler, 1962; Rosenberg and Birdsall, 1986). In some cases the expansion of the early adopters occurred through the destruction of business rivals, in others by merger with them.

Not only did the Prussian administrative system make large, complex organizations relatively efficient, it seemingly made them inevitable. Only very large organizations could fully exploit the Prussian administrative system. Only they could capitalize on extreme task specialization or afford the throngs of staff experts needed to gather and process quantities of data for top management. Hence, for a long time it seemed that bigger organizations were necessarily better. And there seemed to be no natural limits to this conclusion. The planning and control system the General military bureaucracy, under General Erich Ludendorff, used to mobilize Germany’s resources during World War I (the Kriegswirtschaftsplan) was merely an amplification of its peace-time arrangements. The centralized planning system, Gosplan, used in the Soviet Union to implement its long-term policies and strategic plans was an adaptation of the Kriegswirtschaftsplan.

Improvements in information processing, especially in the realms of accounting and finance, eventually limited organizational expansion, however. These innovations had the effect of increasing the relative efficiency of coordinating organizational activities and the flow of materials through arm’s length relationships (as opposed to direct supervision), making it possible to avoid some of the opportunity costs inherent to rule-based governance systems.

The administrative system developed by Alfred Sloan and Donaldson Brown at General Motors in the 1920s demonstrated the maturity of these innovations. Sloan is best known for the multi-product or M-form organizational structure, in which each major operating division serves a distinct product market. Short-run integration under Sloan’s system was achieved via buyer-seller relationships between GM’s five automotive divisions and the divisions making automotive components (e.g., Fisher Body or Delco-Remy). Longer-run integration was achieved via the capital budgeting system devised in 1923 by Donaldson Brown, GM’s chief financial officer.

GM’s operating divisions were managed entirely by the numbers from a tiny corporate headquarters, using the DuPont system of financial control, also devised by Brown. Under this system, each division kept its own books and managers were evaluated in terms of a return-on-assets target. The operating division managers continued to rely on control by rules and standard operating procedures and detailed resource-requirements plans. Sloan, however, believed that it was inappropriate, as well as unnecessary, for top managers at the headquarters level to know much about the details of division operations. If the numbers showed that performance was poor, it was time to change the division manager. Division managers with consistently good numbers got promoted, ultimately to headquarters (Womack, Jones, and Roos, 1990). The divisional form of organization is not only a device to
resolve a span of control problem; it also allows each division or business to be remotely controlled by the numbers from a strategic apex.

The general device that allows for remote control is a control system that aligns the incentives of operating unit managers with the purposes and priorities of the organization as a whole.\(^1\) For a remote control system to operate effectively, financial and cost information need to be relevant. Establishing reporting entities corresponding to segmented business activities is the fundamental rule of thumb to be followed in the construction of such a system. A division is both a reporting entity and a segmented business activity. An ideal type—and even typical—division is one headed by a general manager who reports to the strategic apex and enjoys full line authority over the middle line and operating core.

### Responsibility budgeting

Responsibility budgeting is the most common remote control system used by large-scale organizations in the private sector. It is a form of internal contracting in which: (a) units and managers are evaluated relative to the targets they accept, (b) only financial measures are used to measure and reward accomplishment or punish failure, and (c) financial success or failure is attributed entirely to managerial decisions and/or employee performance. While private businesses were quick to learn bureaucratic control from government, governmental organizations have been slow to adopt remote control systems.

The digression in the last section of this article is relevant here because responsibility budgeting is as much organizational engineering as it is financial management and accounting. Organizational engineering is concerned with the following three elements:

- **Administrative structure**—the structure depicted in an organization chart showing the organization’s administrative units and their relationships to each other. Under responsibility budgeting, work can be arranged into administrative units according to mission, function, and/or region.
- **Responsibility structure**—the allocation of authority and responsibility to individuals within the organization. Under responsibility budgeting, authority and responsibility must be unambiguously assigned.
- **The account or control structure**—the system of measuring and evaluating performance. Under responsibility budgeting information on inputs, costs, activities, and outputs is critically important.

Under a fully developed responsibility budgeting and accounting system, administrative units and responsibility centers are coterminous and fully aligned with the organization’s account structure, since the information it provides can be used to coordinate unit activities as well as to influence the decisions of responsibility center managers.

Under responsibility budgeting, two basic rules govern organizational design. First, organizational strategy should determine structure. Strategy means the pattern of purposes and policies that defines the organization and its missions and that positions it relative to its environment. Single mission organizations should therefore be organized along functional lines; multi-mission organizations should be organized along mission lines; multi-mission,
multifunction organizations should be organized along matrix lines. Where a matrix organization is large enough to justify an extensive division of labor, responsibility centers should be designated as either mission or support centers, with the latter linked to the former by a system of internal markets and prices (transfer pricing).

The second basic rule is that the organization should be as decentralized as possible. Most students of management believe that the effectiveness of large, complex organizations improves when authority and responsibility are delegated down into the organization. Of course, authority should not be delegated arbitrarily or capriciously. Decentralization requires prior clarification of the purpose or function of each administrative unit and responsibility center, procedures for setting objectives and for monitoring and rewarding performance, and an account structure that links each responsibility center to the goals of the organization as a whole.

The biggest difference between government budgets and responsibility budgets is that government budgets tend to be highly detailed spending or resource acquisition plans, which must be scrupulously executed just as they were approved (Thompson and Jones, 1986). In contrast, operating budgets in the private sector are usually sparing of detail, often consisting of no more than a handful of financial targets. As we noted earlier, Sloan of General Motors, one of the fathers of responsibility budgeting, believed it was inappropriate for corporate managers to know the details of responsibility center operations. The notion that responsibility centers should be managed at arm’s length, by the numbers, from a small corporate headquarters, reflects the effort to delegate authority and responsibility down into the organization. As the OECD report, *Budgeting for Results: Perspectives on Public Expenditure Management* (1995), explains, delegation of authority means giving agency managers the maximum feasible authority needed to make their units productive—or, in the alternative, subjecting them to a minimum of constraints. Hence, delegation of authority requires operating budgets to be stripped to the minimum needed to motivate and inspire subordinates. Under responsibility budgeting the ideal operating budget would contain a single number or performance target (e.g., a production quota, a unit cost standard, or a profit or return on investment target) for each administrative unit/Responsibility center.

In responsibility budget formulation, an organization’s policies, the results of all past policy (capital budgeting, see Thompson, 1997) decisions, are converted into financial targets that correspond to the domains of administrative units and their managers (Anthony and Young, 1994:19). In responsibility budget execution, operating performance is monitored and subordinate managers are evaluated and rewarded. Operating performance targets must be expressed in financial terms. This makes it possible to make comparisons across unlike responsibility centers, thereby permitting the relative performance of managers to be evaluated and increasing the motivational efficacy of internal competition. It also has the effect of keeping higher levels of administration ignorant of operating details, thereby discouraging them from meddling in the affairs of their responsibility center managers.

**Types of responsibility centers**

Responsibility centers are usually classified according to two dimensions:
The integration dimension—i.e., the relationship between the responsibility center’s objectives and the overall purposes and policies of the organization; and

The decentralization dimension—i.e., the amount of authority delegated to responsibility managers, measured in terms of their discretion to acquire and use assets.

On the first dimension, a responsibility center can be either a mission center or a support center. The output of a mission center contributes directly to an organization’s objectives or purpose. The output of a support center is an input to another responsibility center in the organization, either another support center or a mission center.

On the decentralization dimension, accountants distinguish among four types of responsibility centers based on the authority delegated to responsibility managers to acquire and use assets. Discretionary expense centers, the governmental norm, are found at one extreme; profit and investment centers are at the other. A support center may be either an expense center or a profit center. If the latter, its profit is the difference between its costs and its “revenue” from “selling” its services to other responsibility centers. (Sells is in quotation marks here because the organization as a whole has not sold anything to an outside party. Rather, the responsibility center providing the service records revenue in its accounts and the center receiving the service records an expense. Revenue and expense cancel out when the organization consolidates its books. Money rarely changes hands in interdivisional transfer pricing. Responsibility centers don’t get to keep “their” profits. Only the organization as a whole earns a profit. Selling to and buying from outsiders are the only activities that can generate real profits or losses for an organization). Both profit and investment centers are usually free to borrow, and investment centers are also free to make decisions about plant and equipment, new products, and other issues that are significant to the long run performance of the organization.

Discretionary expense centers incur costs. The difference between them and other kinds of responsibility centers is that their managers have no independent authority to acquire assets. Instead, the manager’s superiors must authorize each acquisition. In the US system, under detailed line item budgets, acquisitions must be authorized by Congress and signed into law by the President. But all discretionary expense center managers are accountable for compliance with an asset acquisition/resource requirements plan (expense budget), whether written into law or not. Once acquisitions have been authorized, discretionary expense center managers are usually given considerable latitude in their deployment and use. Managerial accountants generally believe that administrative units should be discretionary expense centers only where there is no satisfactory way to match their expenses to final cost objects.

In some cases, expense center managers are evaluated in terms of the number and type of activities performed by their center. Where each of the activities performed by the center earns revenue or is assigned notational revenue (transfer price) by the organization’s controller, these centers are referred to as revenue centers.

In a cost center, the manager is held responsible for producing a stated quantity and/or quality of output at the lowest feasible cost. Someone else within the organization determines the output of a cost center—usually including various quality attributes, especially delivery schedules. Cost center managers are usually free to acquire short-term assets (those that are
wholly consumed within a performance measurement cycle), to hire temporary or contract personnel, and to manage inventories.

In a standard cost center, output levels are determined by requests from other responsibility centers and the manager’s budget for each performance measurement cycle is determined by multiplying actual output by standard cost per unit (see Thompson, 1998). Performance is measured against this figure—the difference between actual costs and standard costs.

In a quasi-profit (or pseudo-profit) center, performance is measured by the difference between the notational revenue earned (transfer price) by the center and its costs (Kaplan and Cooper, 1998: 56–73). For example, let’s say a Veteran’s Administration hospital department of radiology performed 500 chest X-rays and 200 skull X-rays for the department of geriatrics. The notational revenue earned was $25 per chest X-ray (500) = $12,500 and $50 per skull X-ray (200) = $10,000, or $22,500 total. If the radiology department’s costs were $18,000, it would earn a quasi-profit of $4,500 ($22,500 – $18,000).

In profit centers, managers are responsible for both revenues and costs. Profit is the difference between revenue and cost (or expense). Thus, profit center managers are evaluated in terms of both the revenues their centers earn and the costs they incur. In addition to the authority to acquire short-term assets, to hire temporary or contract personnel, and to manage inventories, profit center managers are usually given the authority to make long-term hires, set salary and promotion schedules (subject to organization-wide standards), organize their units, and acquire long-lived assets costing less than some specified amount.

In investment centers, managers are responsible for both profit and the assets used in generating the profit. Thus, an investment center adds more to a manager’s scope of responsibility than does a profit center, just as a profit center involves more than a cost center. Investment center managers are typically evaluated in terms of return on assets (ROA), which is the ratio of profit to assets employed, where the former is expressed as a percentage of the latter. In recent years many have turned to economic value added (EVA), net operating “profit” less an appropriate capital charge, which is a dollar amount rather than a ratio and is more generally consistent with the value-creating purposes of organizations (Kaplan and Cooper, 1998: 265–270; Carlton and Perloff, 1996: 334–341, 77–78, 373–374).

Formerly, in most large complex organizations in the private sector, individual production units were typically standard cost centers; staff units were typically discretionary expense centers. Indeed, only mission centers were allowed to be investment centers. The reasons for this are complex, but they go to difficulties associated with expensing intermediate and joint products. Mission centers in private sector organizations produce final products that are easily priced and that are expensed following generally accepted accounting practice. In contrast, support centers produce intermediate products and these were, until recently, hard to cost, let alone price, with accuracy. Attempts to do so were often either excessively arbitrary or prohibitively costly.

Nowadays, however, advances in information technology, managerial accounting, and organizational design have made it possible and, in some cases, beneficial to treat every responsibility center in an organization as an investment center (see Thompson, 1998).

Paradoxically, public sector organizations are a mirror image of large complex organizations in the private sector. We know now how to treat their support centers as quasi-profit or
even investment centers (Lapsley, 1994; for additional public sector examples, see Anthony and Young, 1994: 371–374; Kaplan and Cooper, 1998: 245–251). But, because the final products of government’s core mission centers are public goods that are passively enjoyed (Vining and Weimer, 1998), pricing final outputs remains for the time being and for the foreseeable future either excessively arbitrary or prohibitively costly. This means, for example, that while it might make sense to treat military depot maintenance, spare parts management, or facilities support centers as investment centers, it will continue to be necessary to treat the armed forces’ combatant commands as discretionary expense centers. Fortunately, as far as exhaustive expenditures are concerned, about 75 percent of the activities performed by the US federal government fall into the support category and, for the most part, state and local governments are not in the business of supplying pure public goods (see Goldin, 1977).

Transfer pricing

Under responsibility budgeting, support centers provide services or intermediate goods to other responsibility centers in return for a notational transfer price, organizations are structured to take advantage of specialized knowledge and local conditions, and center managers make decisions and are held responsible for the overall financial performance of their centers. Sound transfer pricing is, therefore, the key to aligning the incentives of responsibility center managers with organizational interests.

Transfer pricing is also important to transparency within organizations. It helps to determine the costs of services provided by one unit to another, which is central to measuring performance relative to a financial target, and therefore plays a major role in establishing, as well as manipulating, the incentives facing responsibility center managers. Transfer pricing also reveals the internal costs of service decentralization where costs are incurred in transferring decision rights to others within an organization. When one sub-unit transfers tangible assets, knowledge, skills, etc., to another, both units calculate the cost as a means of revealing their liquid and tangible asset use internally and in external provision of service.

There are two common approaches to transfer pricing:

- Laissez-faire transfer pricing: buying and selling responsibility centers are completely free to negotiate prices, to deal, or not to deal; and
- Marginal or incremental cost pricing: the responsibility center selling the service is required to charge the buying responsibility center whichever is less of market or incremental cost.

(A third method is based upon fully distributed average cost of the service or product.)

However, the circumstances that justify large complex organizations—economies of scale and scope—render these simple transfer-pricing mechanisms problematic. Scale economies are usually the result of large, lumpy investments in specialized resources—technological knowledge, product specific research and development, or equipment. These investments tend to give rise to bilateral monopoly, a circumstance that provides an ideal environment for opportunistic behavior on the part of suppliers and customers. For example, once an intermediate product producer has acquired a specialized asset, customers may be able to
extract discounts by threatening to switch suppliers. In that case, the supplier may find it necessary to write off a large part of the specialized investment. Or, if demand for the final good increases greatly, the intermediate product supplier may be able to extort exorbitant prices from customers. Hence, where the relationship between intermediate product supplier and customer is at arm’s length, opportunistic behavior may eliminate the payoff to what would otherwise be cost effective investments. For example, the Report of the Commission on Roles and Missions of the Armed Forces (Commission, 1995; see also Thompson and Jones, 1994) suggested that budget authority should flow through the combatant commands to the military departments. Were that the case, lacking a long-term credible commitment on the part of the Joint Chiefs and the combatant commanders, the navy’s investment in specialized assets like aircraft carriers would permit it to be exploited in peacetime. In wartime, of course, the tables would be turned.

The new economics of organizations tells us that vertical integration occurs because it can mitigate this problem, in part through the substitution of direct supervision for remote control (see Williamson, 1985). For example, in a study of military procurement, Scott Masten (1984) demonstrated that specialized investments are critical to vertical integration. Where intermediate products were both complex and highly specialized (used only by the buyer), there was a 92 percent probability that they would be produced internally; even 31 percent of all simple, specialized components were produced internally. The probability dropped to less than 2 percent if the component was unspecialized, regardless of its complexity.

Unfortunately, the problems that arise in arm’s length transactions where there are few alternative suppliers/customers also arise where one attempts to replicate free market forces within the organization, allowing buying and selling responsibility centers complete freedom to negotiate prices (laissez-faire transfer pricing). Traditionally, economists have argued that services should be transferred at marginal or incremental cost to the buying responsibility center. But this can seriously distort the evaluation of support center performance and tend to eliminate incentives to improvement.

As a result, organizations face a serious dilemma. They can maximize short-run performance by using marginal cost in internal transactions, thereby seriously distorting performance measurement and incentives, which will cause shortfalls in long-run performance. Or they can sacrifice short-term performance by relying on laissez-faire transfer pricing, thereby obtaining superior measures of the support center’s contributions to organizational performance, and improve the chances of maximizing performance in the long term. Organizations can promote short-run performance by using incremental cost pricing or they can promote long-term performance by using laissez-faire pricing, but they cannot do both simultaneously using either of these simple transfer pricing mechanisms.

In theory, bilateral monopoly can be governed quite satisfactorily by unbalanced transfer prices, multi-part transfer prices, or quasi-vertical integration. Under unbalanced transfer prices, the selling responsibility center is credited with the full cost of the transacted item (often standard cost) plus an agreed upon markup, the buying center is charged its marginal cost, and the organization’s accounts are adjusted to reflect the difference between the two. Unbalanced transfer prices are rarely used, however, where market prices are available. Under multi-part transfer prices, the service delivered is decomposed to reflect underlying cost drivers and priced accordingly (your home phone bill is an excellent example of a
multi-part tariff). Under quasi-vertical integration, the buyer invests in specialized resources
and loans, leases, or rents them to their suppliers. Quasi-vertical integration is common in
both the automobile and the aerospace industries, and, of course, it is standard procedure for
the Department of Defense to provide and own the equipment, dies, and designs that defense
firms use to supply it with weapons systems and the like (see Monteverde and Teece, 1982).
Other organizations that rely on a small number of suppliers or a small number of distributors
write contracts that constrain the opportunistic behavior of those with whom they deal.

In still other cases, desired outcomes can be realized through alliances based on the
exchange of hostages (e.g., surety bonds, exchange of debt or equity positions) or just plain
old-fashioned trust based on long-term mutual dependence. Toyota, for example, relies on a
few suppliers that it nurtures and supports (Womack, Jones, and Roos, 1990). They have
substantial cross-holdings in each other and Toyota often acts as its suppliers’ banker. Toyota
maintains tight working links between its manufacturing and engineering departments and its
suppliers, intimately involving them in all aspects of product design and manufacture. Indeed,
it often lends them personnel to deal with production surges and its suppliers accept
Toyota people into their personnel systems.

Toyota’s suppliers are not completely independent companies with only a marketplace
relationship to each other. In a very real sense, they all share a common purpose and destiny.
Yet Toyota has not integrated its suppliers into a single, large bureaucracy. It wanted its
suppliers to remain independent companies with completely separate books—real profit/
investment centers, rather than merely notational ones—selling to others whenever possible.
Toyota’s solution to the bilateral monopoly problem appears to work just fine (Womack,
Jones, and Roos, 1990). In fact, with the exception of unbalanced transfer prices, none of the
solutions to the bilateral monopoly problem noted here presumes vertical integration. All that
is required is full access to cost and production information (Milgrom and Roberts, 1992).
Of course, all of these solutions to the transfer pricing/organizational design are potentially
available to government organizations. Indeed, many of them were pioneered by federal
acquisitions personnel or imposed by public utility commissions. They are not, however,
widely understood or appreciated by public administrators and financial managers.

Responsibility budgeting in government

The origins of responsibility budgeting and accounting in government can be traced to the
Planning, Programming, and Budgeting System (PPBS) era in the US Department of Defense
(1961–1967). Responsibility budgeting and accounting was the centerpiece of Project Prime,
perhaps the most promising of the organizational design and development efforts initiated
under Secretary of Defense Robert McNamara. Project Prime was the brainchild of Robert
N. Anthony (Juola, 1993: 43–44), who succeeded Charles Hitch as defense controller in
September 1965. Anthony saw the need for clarification of the purpose of each of the
administrative units that comprised the Department of Defense, their boundaries, and their
relationships to each other, and for an account structure that would tie the entire organization
together. Anthony (1962) proposed that the Department of Defense:
- Classify all administrative units as either mission or support centers;
- Charge all costs accrued by support centers—including charges for the use of capital assets and inventory depletion—to the mission centers they serve;
- Fund mission centers to cover their expected expenses—including support center charges;
- Establish working capital funds to provide short-term financing for support units; and
- Establish a capital asset fund to provide long-term financing of capital assets and to encourage efficient management of their acquisition, use, and disposition.

The principal formal device by which a measure of intra-organizational decentralization was and is accomplished within the US Department of Defense is the revolving fund. These funds involve buyer-seller arrangements internal to the Department of Defense. They have actually been in use for some time. The navy had a revolving fund as early as 1878. Modern-day revolving funds date to the 1947 National Security Act, which authorized the defense secretary to use them to manage support activities within the Department of Defense. Two kinds of funds have been established under this authority: stock and industrial funds. Stock funds are used to purchase supplies in bulk from commercial sources and hold them in inventory until they are supplied to the customer—usually a military unit or facility. Industrial funds are used to purchase industrial or commercial services (e.g., depot maintenance, transportation, etc.) from production units within the Department of Defense. Both kinds of funds are supposed to be financed by reimbursements from customers’ appropriations (Juola, 1993:43).

Anthony’s proposal would have expanded the scope of this device and enhanced its effectiveness by establishing rules for setting transfer prices prospectively rather than retrospectively and by making support center managers responsible for meeting explicit financial targets. Internal buyer-seller arrangements encourage efficient choice on the part of support centers, as well as the units that use their services, only if prices are set ahead of time and support centers charge all of their costs against revenues earned delivering services. Furthermore, their managers must be fully authorized to incur expenses to deliver services, and held responsible for meeting the stated financial goals of their centers (Bailey, 1967:343).

Project Prime failed. One reason for its failure is that the federal government of the US accounts for purchases, outlays, and obligations, but it still does not account for consumption. Full value from the application of responsibility budgeting can be obtained only where government adopts a meaningful form of consumption or accrual accounting (measuring the cost of the assets actually consumed producing goods or services). Because the US government does not account for resource consumption, its cost figures are necessarily statistical in nature (i.e., they are not tied to its basic debit and credit bookkeeping/accounting records). Without the discipline that debit and credit provides, these figures are likely to be satisfactory only for illustrative purposes or where a decision maker must make a specific decision and a cost model has been tailored to the decision maker’s needs. Another reason for the failure of Project Prime is that the US appropriations process does not perform the capital budgeting function satisfactorily, a problem that PPBS did not really address and certainly didn’t fix.
Besides which, the existing process procrusteanizes every operating cycle to fit the fiscal year. Responsibility budgeting next surfaced in the United Kingdom, as part of the Thatcher government’s Financial Management Initiative, announced May 17, 1982 (Pollitt, 1993; Lapsley, 1994). The Financial Management Initiative called for a radical change in the internal structure and operations of government agencies. Objectives were to be assigned to responsibility centers. Costs were to be systematically identified. They were to be measured on an accrual basis (i.e., matching resources consumed to services delivered) and include not only the direct costs of service delivery but overheads as well. This identification enabled those responsible for meeting particular objectives to be held accountable for the cost of the resources they were consuming.

The scope of responsibility accounting and budgeting in the UK was further extended in 1988 by the Thatcher government’s Next Steps Initiative. In the last eight years, much of the British civil service has been reorganized into a set of executive agencies that have been given considerable administrative and fiscal flexibility and expected to meet annual financial performance targets. The heads of these executive agencies are no longer career civil servants. They are recruited from either the private sector (about 25 percent) or public sector, hired on short term contracts, with pay and tenure contingent on their success in meeting annual performance targets. By April 1996 there were 125 executive agencies in the UK, with 37 more candidates under consideration, covering about 75 percent of the British civil service (Roberts, 1997).

Following the launch of the Financial Management Initiative in Great Britain, other governments—Australia, Canada, Denmark, Finland, and Sweden—have adopted responsibility budgeting and accounting. None, however, has moved as far or as fast as New Zealand. Moreover, New Zealand’s reformers explicitly recognized their debt to agency theory (Boston et al., 1996; Lewis et al., 1996).

New Zealand

Most of the external attention given to New Zealand’s public management reforms has focused on its efforts to improve transparency: the adoption of accrual accounting and reporting on performance. New Zealand was the first country to publish a full set of standard financial statements, including a balance sheet of assets and liabilities and an accrual based operating statement of income and expenses. However, the changes made in the structure of the government of New Zealand designed to promote effective resource use and investment are even more significant than are the changes in financial reporting. First of all, New Zealand’s Parliament privatized everything that was not part of the core public sector. The residual core public sector now includes a mix of policy and regulatory and operational functions and the military services, policing and justice services, social services such as health, education, and the administration of benefit payments, research and development, property assessment, and some other financial services.

Second, Parliament redefined the relationship between it and the heads of government agencies. Agency heads lost their permanent tenure and are now known generically as ‘chief executives.’ They are appointed for fixed terms of up to five years, with the possibility of
reappointment. Each works to a specific contract, the conditions of which are negotiated with
the State Services Commission and approved by the Prime Minister. The State Services
Commission also monitors and assesses executive performance. Remuneration levels are
directly tied to performance assessment.

Third, Parliament changed the way it appropriates funds for use by the remaining
government agencies to link appropriations to performance, allowing Parliament fiscal
control, but, at the same time, providing greater fiscal flexibility for agency heads. The basis
of appropriation depends on the agency’s ability to supply adequate information about its
performance. Three modes of appropriation are possible, recognizing that some agencies
provide goods and services that are more commercial or contestable than do others.\(^9\)

All agencies started out in Mode A, but most have progressed to Mode B and a few to
Mode C. Under Mode A, agencies were discretionary expense centers and Parliament
appropriated funds for the purchase of resources. Indeed, the only change from the budget
process in effect before 1989 (or, for that matter, the budgets used by most governments
throughout the world) is that separate appropriations were provided for expenditures for plant
and equipment. This mode remained in force until the agency developed a satisfactory
accrual accounting system and identified its outputs, both of which are needed for perfor-
mance assessment.

Under Mode B, most agencies are quasi-profit centers. This mode is designed for agencies
that supply traditional, noncontestable, governmental services: the central control agencies,
including the State Services Commission, most regulatory and police functions, and some
justice services, i.e., policy agencies and activities that include an element of compulsion for
the buyer. Under this mode, Parliament appropriates funds retrospectively to reimburse
agencies for expenses incurred in producing outputs during the period covered by the
contract, whether for the government or third parties. Costs are measured on an accrual basis;
they include depreciation, but exclude taxes and the return on funds employed. Changes in
an agency’s net asset holdings are also explicitly appropriated.

Under Mode C, agencies are investment centers. Appropriations pay for the outputs
produced by the agency and for any changes in the agency’s net assets. Agencies in Mode
C are required to pay interest, taxes, and dividends and must establish a capital structure.
Mode C agencies are set up in a competitively neutral manner so that their EVAs can be
assessed by comparison with firms in the private sector. The prices paid for the outputs
supplied by Mode C agencies are supposed to approximate fair market prices. In general, this
means that agencies must show that they are receiving no more than the next best alternative
supplier would receive for providing the outputs. Mode C agencies are not permitted to
borrow on their own behalf or to invest outside their own areas of operation. Each month,
each agency reports on its financial position, cash flow, resource usage, and revenue by
output. Variances are calculated and explanations provided. Under Mode B, managers are
free to make some decisions about investments in plant and equipment. Managers could
make even more decisions under C. The fact that their financial performance is one of the
main bases upon which managerial performance is assessed helps insure that those decisions
are sound. Unfortunately, far too few of New Zealand’s agencies have actually converted to
Mode C (see Schick, 1996; for criticisms of the New Zealand reforms, see IPMJ, Vol. 3, No.
1, Mascarenhas, 1996, and Quiggen, 1998).
Government’s key decisions remain firmly in the hands of Parliament. The decisions that have the most significant future consequences for the government of New Zealand’s stakeholders are clearly those which have to do with the kind, quantity, and quality of service provided by the citizenry. Under the existing system of appropriations and financial reporting, those issues must be explicitly confronted when the cabinet enters into long-term contracts with agencies, state owned enterprises, and firms to deliver service outputs, and its consequent liabilities must be stated in present value terms.

The United States

Responsibility budgeting and accounting was adumbrated in the United States and influenced the now defunct Defense Management Report Initiatives of the Bush/Cheney era in the Department of Defense, and arguably the content of both the Chief Financial Officers Act and the National Performance Review’s calls for mission-driven, results-oriented budgets and, more recently, performance based organizations (OECD, 1995: 230). Still, it has had little or no practical effect in this country.

There are two explanations for this fact. The first is that many students of the expenditure process reject the notion that remote control can be reconciled with the American legislative budgetary process. Some people even assert that it can be practiced only by responsible unitary governments on the Westminster model, although that claim is belied by the Swiss and Swedish examples (Schedler, 1995; Arwidi and Samuelson, 1993) and various state (Barzelay, 1994) and local governments here in the US (Kaplan and Cooper, 1998: 245–251). Of course, it would not be easy to reconcile responsibility budgeting with the American legislative process, but we do not believe that they are necessarily incompatible (see Thompson, 1994; Harr, 1989; Harr and Godfrey, 1991, 1992). A second possible explanation for its failure to leave its mark on government accounting and budget practices in the United States is that, unlike most other countries, America has large, well-organized corps of government accountants, auditors, budgeters, program analysts, and teachers of government accounting and budgeting. All of these groups have vested interests in differentiating public from private practice, because that difference gives value to their expertise. A third reason seems to be that many people, in and outside of government, evidently believe that “public” necessarily implies Prussian-style bureaucracy. Where the purpose of the organization in question or the technology available to it make Prussian-style bureaucracy inappropriate, they will hear of no alternative short of full-scale privatization.

The Clinton Administration’s second-term government reform efforts have centered on pushing the concept of performance-based organizations (PBOs) modeled after Britain’s Financial Management Initiative (Green, Jones, and Thompson, 2000; see, however, Roberts, 1997). The main theme of this reform effort is the use of contracts to hold PBOs accountable for financial performance.

The progress of this effort has been glacial. When legislation for the first PBO candidate, the Patent and Trademark Office, was sent to Congress, it aroused an intense debate between the administration and the chair of the House Judiciary subcommittee on courts and intellectual property regarding the relative merits of the PBO model versus a corporate model. This debate has been reproduced in various venues for successive PBO candidates.
What goes around, comes around

It is somewhat ironic that governments are beginning to embrace remote control at the same time many well-managed businesses are abandoning it (Bruggeman, 1995; Otley, 1994; Bunce, Fraser, Woodcock, 1995). These businesses have abandoned remote control because they are no longer compartmentalized the way they once were and it simply doesn’t reflect the way they are now put together (Bruggeman, 1995; Otley, 1994; Bunce, Fraser, and Woodcock, 1995). Arguably, decompartmentalization is being driven by the information revolution, which is breaking down economies of scale and scope built upon functional specialization (Reschenthaler and Thompson, 1996). According to Michael Hammer, modern data bases, expert systems, and telecommunications networks provide many, if not all, of the benefits that once made internal specialization of administrative functions like personnel, finance, accounting, etc. attractive (Hammer, 1990: 108–112). To the extent that the provision of these services requires specialized skills, they are increasingly contracted out to specialist firms. The people in the organization who actually do its real work perform the rest.

Hammer claims that jobs should be designed around an objective or outcome instead of a single function; that functional specialization and sequential execution are inherently inimical to expeditious processing; that those who use the output of activity should perform the activity and the people who produce information should process it, since they have the greatest need for information and the greatest interest in its accuracy; that information should be captured once and at the source; that parallel activities should be coordinated during their performance, not after they are completed; and last, that the people who do the work should be responsible for decision making and control built into job designs (Hammer, 1990).

Decompartmentalization has led to smaller, flatter organizations, organized around a set of generic value-creating processes and specific competencies. Some single-mission organizations are now organized as virtual networks, some multi-mission organizations as alliances of networks. Philip Evans and Thomas Wurster refer to both of these kinds of organizational arrangements as hyperarchies, after the hyperlinks of the World Wide Web (Evans and Wurster, 1997: 75). Evans and Wurster assert that these kinds of organizations, like the Internet itself, the architectures of object-oriented software programming, and packet switching in telecommunications, have eliminated the need to channel information, thereby eliminating the tradeoff between information bandwidth (richness) and connectivity (reach). Evans and Wurster describe virtual networks (structures designed around fluid, team-based collaboration within the organization) as deconstructed value chains, and alliances of networks (the pattern of “amorphous and permeable corporate boundaries characteristic of companies in the Silicon Valley”) as deconstructed supply chains, in which “everyone communicates richly with everyone else on the basis of shared standards.”

The system used by IBM at its plant in Dallas, Texas, is an example of an existing virtual network. It has been designed to mimic a market-like, self-organizing system. Everyone in the organization plays the part of customer or provider, depending on the transaction, and the entire plant has been transformed into a network of dyads and exchanges. Each exchange is a closed loop involving four distinct steps: request from a customer and offer from a provider, negotiation of the task to be performed and the definition of success, performance,
and customer acceptance. Until this last step is completed, the task remains unfinished. Each closed loop of workflow is further broken down into subloops. Under this system, even simple tasks give rise to dozens of loops and interconnecting lines; more complex tasks, such as modifying a major product, to hundreds; and managing the entire Austin plant to thousands. IBM uses powerful computers to keep track of all of these loops and lines, to chart all activities and operational flows within the plant, to keep track of progress being made at each stage of each transaction, and to prod tardy participants into action—this is control built into job design with a vengeance.

The effect of this system has been to break down departmental boundaries, eliminate bottlenecks, and to empower employees to take initiatives and coordinate themselves. As a byproduct, the computer systems that keep track of all these loops and lines also identify the resources going into a particular job, almost entirely eliminating the need for cost allocation. Moreover, this information is available both prospectively and retrospectively to anyone in the organization.

Some well-managed multi-mission organizations such as Johnson & Johnson, 3M, and Rubbermaid have already organized themselves into loose alliances of networks, sharing only their top management, a set of core competencies, and a common culture (Quinn, 1992). The control systems used by these organizations are like those of centralized bureaucracies in that they collect a lot of real-time information on every aspect of operations, including nonfinancial information (see Table 1), but unlike the control systems of stovepiped centralized bureaucracies, which were erected on the premise that the exercise of judgment should be passed up the managerial ranks, this information is used to push the exercise of judgment down into the organization, to wherever it is needed, at the point of sale, at delivery, or in production (Simons, 1995). From top management’s perspective, the primary purpose of this information is to provide them with insight into the integrity, competence, and morale of their network managers and employees so that they can allocate their best people to the most important jobs.

How far hyperarchy will go is an open question. Evans and Wuster (1997) claim that it will destroy all hierarchies, whether of logic or of power, “with the possibility (or the threat) of random access and information symmetry.”

If hyperarchy is where we are all heading, responsibility budgeting and accounting is at best an intermediate stage (Otley, Broadbent, and Berry, 1995). It is now apparent, as it really was not before, that responsibility budgeting restricts the upward flow of operating information within organizations—making decentralization a necessity as well as an ideal. In
contrast, networks and alliances are information rich environments. For the most part, access to information is symmetrical in fully networked organizations—equally available to all the people in the organization.

Why not skip the intermediate stages and go directly to networked organizations? One of the referees for this article suggested that some governments might be moving in that direction. He notes the trend toward performance efforts and accomplishments reporting and the widespread acquisition of so-called enterprise resource planning (ERP) systems built around common data structures and centralized information warehouses, which permit data to be entered and accessed from anywhere in the organization (SAP, PeopleSoft, Oracle, etc.)

Moreover, as Martha Feldman and Anne Khademian (this issue) correctly observe, both traditional bureaucratic controls and responsibility budgeting and accounting are inherently hierarchical systems. Neither comport well with democracy. One can envision, instead, fully networked organizations embedded in a web oriented, wireless system of governance, utilizing a wide range of “knowledge management” tools, digital agents, and (group) decision support systems” to realize the ideal of full democratic participation (6, 2001).

These tools include:

- Systems dictionaries of cultural or disciplinary dialects that enable participants with different backgrounds to understand one another’s vocabulary;
- Mental map and mental representation tools that enable participants to develop graphical representations of their own or others’ basic conceptual approach to problems;
- Document profiling systems in shared work spaces that enable participants to hyperlink related documents, to identify key relationships with key documents, etc.;
- Memory capture systems designed to track and enable retrieval in a variety of forms, styles, practices, precedents, contacts, cultures, that may be relevant to future decision making;
- Open hypermedia data systems that enable new participants quickly to develop an understanding of the community’s prior history, culture, distinctive uses of terms, etc.;
- Learning models, including neural net based models of professional judgment and rule-based expert systems;
- Individual and group creativity tools including idea generation tools, electronic whiteboards for graphical representation of connections between ideas;
- Graphical problem structuring tools that are integrated with search and analysis agents operating over banks of relevant information available on intranets, perhaps organized using hypermedia linkages; and
- Meeting management tools to help participants generate options, identify pros and cons, and track the flow of discussion and debate.

These knowledge management systems foretell the transformation of collective decision making. They are potentially both richer than paper and ink systems and synchronous. They might allow participants to work simultaneously rather than sequentially. This means that knowledge management systems would permit effective 360-degree sharing and governing. Hence, they could enable all the participants in the polity to shape ideas, problems, arguments, and solutions.

Sounds good! Nevertheless, Robert Kaplan and Robin Cooper (1998: 25) argue that
organizations that try to move directly to a system where everyone communicates richly with everyone else on the basis of shared standards, without passing through a recommended period of experimentation with operational-feedback and cost-measurement systems, will almost surely fail.

We are less certain that this is the case. We acknowledge, however, that organizational decentralization can work in an information-rich environment only where top management attends to top management functions—strategic planning, organizing, staffing, the intellectual and cultural development of the organization—and refrains from meddling in the conduct of operations. This takes self-restraint, and self-restraint must be learned. For that reason, it may make sense for governments to experiment with responsibility budgeting rather than going directly to new modes of organization and control. Besides, the brave new word of hyperarchy is not yet here and perhaps not even imminent.

Notes

1. Of course, traditional bureaucratic administrative arrangements are also implicitly contractual devices that seek to align the interests of agents with those of their principals. For a discussion of how this works under discretionary expense budgets, see Thompson, 1993; under cost-plus contracts, see Thompson and Jones, 1986.

2. We recognize that some operating flexibility is built into the discretionary expense budgets that are enacted into law—e.g., journal transfers and deficiency mechanisms—but this flexibility is very narrowly circumscribed and also highly rule governed (see Pitsvada, 1983).

3. This section is based on the discussion in Anthony and Young, 1994.

4. This is a very important practical point. In the US federal government revolving fund, agencies are prohibited from earning a “profit.” Instead, they are generally directed to operate on a breakeven basis. The motivational effect of this maximand is to cause their actual costs to exceed standard cost in the majority of instances. In contrast, if they were directed to maximize “profit” and, assuming that they continued to base per-unit user charges on historical fully distributed average costs, the effect would be to encourage them to save budget authority for their internal customers and dollars for the US Treasury. It might also have the effect of ratcheting their unit costs down, instead of permitting them to creep gradually up as is now typically the case. Clearly, the breakeven policy doesn’t make sense and probably reflects the failure to understand the simple point that selling to and buying from outsiders are the only activities that can generate real profits or losses for the organization. Interestingly, many revolving fund agencies are permitted to earn a profit for their parent organization when they sell outside the organization (e.g., Air Force product support for foreign military sales).

5. These are in fact currently revolving-fund operations; they tend to be treated like revenue centers, except when they “lose money,” which is often the case.

6. When factors enter into joint production, they typically develop a degree of specificity with respect to each other. Specificity gives rise to a Williamsonian “Fundamental
Transformation” from an *ex ante* competitive relationship to an *ex post* bilateral monopoly (see Williamson, 1985; Milgrom and Roberts, 1992).

7. OMB’s Justine Rodriguez (1996) would fix this problem by creating a new set of accounts along the lines of the fund accounting systems used by nonprofit schools and hospitals. For example, each department could have one or more capital asset acquisition accounts. Outlays to acquire capital assets would be charged to these accounts, which would hold assets, but perform no operations. These accounts would also be permitted to borrow from the Treasury to acquire assets. The assets they held would be rented/leased to programs, so each program account would show the cost of using assets, but this rent would net out of department totals because of offsetting collections to capital acquisition accounts. In cases where large inventories were acquired, they could be held by intra-governmental support revolting funds (e.g., franchise or working capital accounts) and “sold” just-in-time to programs. Employee pension funds already receive accrual payments from departments. Retiree health benefits could be treated the same way. Similarly, Rodriguez argues that we could require clean-up liabilities be paid to an account that would finance future environmental restoration. To connect resources with results, program budget accounts would be aligned with programs providing goods, services, and transfers to the public. Support budget accounts (e.g., for personnel, legal, and computer services) would be financed by intra-governmental support revolting funds. Under this system, nearly all resources, except perhaps those to the agency head for policy coordination, would go to programs, which would buy their support competitively from their own department, from other departments, or from the private sector. Program outlays would then approximate program costs and could then be fairly related to program outputs.

8. GPRA and the CFO Act, when successfully implemented, would fix two of these problems (Jones and McCaffrey, 1998). The third would require changes in budget law and executive orders (see Thompson, 1994).

9. The following is based on Scott, Bushnell, and Sallee, 1990.

References


