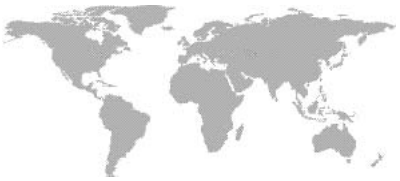
WORLD BANK OPERATIONS EVALUATION DEPARTMENT



Evaluating Public Sector Reform

Guidelines for Assessing Country-Level impact of Structural Reform and Capacity Building in the Public Sector

Navin Girishankar

2001
The World Bank
Washington

http://www.worldbank.org/html/oed

Operations Evaluation Department
Partnerships & Knowledge Programs (OEDPK)
Email: ecampbellpage@worldbank.org
Email: eline@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125

The study was published in the Partnerships and Knowledge Group (OEDPK) by the Ourtreach and Dissemination Unit.

Director-General, Operations Evaluation: Robert Picciotto
Director, Operations Evaluation Department: Gregory Ingram
Manager ; Ruben Lamdany
Task Manager: Navin Girishankar

Abbreviations and Acronyms

AFR	African Region
APL	Adaptable Program Loan
CAE	Country Assistance Evaluation
CAS	Country Assistance Strategy
CDD	Community-Driven Development
CSR	Civil Service Reform
EAP	East Asia and the Pacific Region
ECA	Europe and Central Asia
ESW	Economic Sector Work
ID	Institutional Development
LIL	Learning and Innovation Loan
LAC	Latin America and the Caribbean
MDB	Multilateral Development Bank
M&E	Monitoring and Evaluation
MNA	Middle East and North Africa
MTEF	Medium-Term Expenditure Framework
NGO	Non-Governmental Organizations
NIE	New Institutional Economics
NPM	New Public Management
OED	Operations Evaluation Department
O&M	Operations and Maintenance
PAR	Performance Audit Report
PBD	Planning and Budgeting Department
PER	Public Expenditure Review
PIU	Project Implementation Unit
PREM	Poverty Reduction and Economic Management
PRSC	Poverty Reduction Support Credit
P-SAL	Programmatic Structural Adjustment Loan
SAL	Structural Adjustment Loan
SAR	Staff Appraisal Report
SAS	South Asia
SECAL	Sectoral Adjustment Loan
SSP	Sector Strategy Paper
TAL	Technical Assistance Loan
WDR	World Development Report
WBI	World Bank Institute

Contents

Tables

Figures

Acknowledgments

The paper is part of a larger OED effort to develop methodologies for evaluating the country-level impact of Bank assistance. The author wishes to thank Robert Picciotto and Ruben Lamdany especially for highlighting the centrality of institutional development as an evaluative concern for the Bank and for providing guidance in conceptualizing public management issues in the context of OED's Country Assistance Evaluation product. Many of the ideas presented in the paper grew out of invaluable discussions with Brian Levy (AFTI2) and Nick Manning (PRMPS), both of whom served as peer reviewers. John Johnson, Sohail Malik, and Anwar Shah (OEDCR) provided insightful feedback on issues of CAE methodology, public management, and institutional change. Several others in OED helped shape various iterations of the paper, including Osvaldo Feinstein, Bill Battaile, Rene Vandendries, Poonam Gupta, Timothy Johnston, Laura Kullenburg, Patrick Grasso, and Warren Van Wicklin.

WB16185
C:\DOCUME~1\wb16185\LOCALS~1\Temp\Navins' main paper.DOC
September 24, 2001 4:36 PM

1. Introduction

1.01 Recent sectoral and country reviews of World Bank assistance by the Operations Evaluation Department (OED) have concluded that support for institutional development (ID) efforts—in both sector-specific and public sector management portfolios—has been largely ineffective. Such efforts to improve public management systems were compromised *inter alia* by overly technocratic approaches to institutional design, a bias toward supplying capacity inputs (such as training and equipment) before reforming governance structures, as well as reliance on lending instruments that were not sufficiently flexible to accommodate the complex dynamics of institutional change.[1] In recognition of this sobering track record, the Bank and its partners have begun to rethink the analytical and operational framework that underpins donor-assisted efforts to enhance countries' institutional endowments. While still a work in progress, a consensus approach is emerging, the broad contours of which include greater emphasis on rigorous measurement of institutional performance, use of more flexible programmatic lending instruments, and sequencing strategies that ensure that ID efforts "lock in" improvements in the way public management systems work.[2]

1.02 Even as this consensus approach gains momentum, the development community faces growing pressure to "keep score," not only on the impact of punctual ID efforts by individual donors but also on the overall quality of public management systems in client countries. Scorekeepers or development evaluators will need to improve on existing methodologies for assessing the *relevance* and *efficacy* of ID interventions at the project, sector, and country levels. This will require developing cost-effective techniques for measuring how public management systems affect growth, asset accumulation among the poor, literacy, health status, and other aspects of wellbeing. As part of OED's larger effort to refine its existing methodologies for assessing ID impact, this paper seeks to develop an approach for assessing the *country-level* impact of Bank support for public management systems in client countries.

The Problematique

1.03 Historically, projectized assistance—admittedly limited in terms of flexibility, scope, and scale—was more likely to focus on the finite infrastructural or training needs of beneficiary agencies rather than systemic reform of a particular sector. OED projects followed suit and focused largely on whether projects met their stated objectives rather than capturing the intended and unintended impact on the sector or country as a whole. In recent years, the Bank has sought, through its Country Assistance Strategy (CAS), to channel lending operations as well as advisory services in ways that maximize the country-level impact of its assistance. With country impact as the primary focus of corporate concern, OED evaluations have also "moved to a higher plane (beyond traditional project-level audits)." Specifically, the Country Assistance Evaluation (CAE)—the evaluative analog to the CAS—is a key pillar around which OED has sought to reconstruct the Bank's evaluation architecture.

1.04 Moving to a higher plane—that is, using the country rather than the project as the unit of evaluative analysis—is not without its fair share of problems such as those of aggregation and attribution. These complicate the identification of suitable counterfactuals to help evaluators judge whether Bank involvement actually did make a difference in a particular country. Evaluations of ID include the ancillary challenge of defining "institutions" and then measuring institutional change with a reasonable degree of precision.[3] To evaluate the *country-level* impact

[1] See, for example, OED reviews of the HNP and CSR portfolios
[2] World Bank. *Reforming Public Institutions and Strengthening Governance.* November, 2000.
[3] Goodin, 1996.

of Bank assistance in public management credibly, OED requires (i) a rigorous, yet practical, definition of institutional change, (ii) a robust theory (or set of hypotheses) about the determinants of institutional performance, and (iii) an evaluative methodology to capture the intended and unintended impact of interventions on the overall quality of public management systems.

1.05 *Definition:* Our ability to monitor institutional equilibria and change processes depends in large part on how we define "institutions" in the first place. Numerous definitions of "institutions" currently compete for the evaluator's attention. Historically, OED has used a definition of ID impact in project-level evaluations that is, in theory, broad enough to capture structural, organizational, as well as some cognitive aspects of institutional performance.[4] A metric for country-level impact should capture these various aspects of institutional performance across a range of public management functions such as policymaking, budgeting, and regulation. Equally important is the need to apply standard benchmarks to measure the quality of public management systems across countries with a reasonable degree of uniformity.

1.06 *Attribution:* In addition to benchmarking the performance of public management systems, evaluators should be prepared to attribute changes in performance to endogenous and exogenous factors such as donor-financed interventions, technological change, political dynamics, external shocks, even the "accidents of history."[5] Attributions, to be credible, should derive from hypotheses that are based on a robust theory of institutional change. Unfortunately, a formal theory of institutional change has eluded scholars for the greater part of the twentieth century, as theoretical work in economics and political science all but ignored the constraints that institutions impose on rational actors.[6] While other disciplines such as anthropology were concerned with the role of institutions in social change, this work tended to be "essayist" and offered little by way of formal analysis which could be applied to pressing public policy concerns.[7] As a result, the development literature on institutions was often imprecise, conjectural, and, at times, hortatory.

1.07 In recent years, some economists and political scientists have sought to address the problem of institutional change by formally modeling the constraints that formal and informal rules impose on the preferences of rational actors. Specifically, they sought to explain punctuated equilibria as well as change dynamics in market as well as non-market institutions (for example, legislatures, bureaucracies, local government), using constrained optimization and game theoretic models.[8] A growing body of empirical work using these tools suggests that the "new institutionalism" could provide the basis for integrating the study of market, polity, and community under the rubric of rational choice. Potential applications to development policy include improved diagnosis of institutional constraints, better design of projects and programs, and more rigorous evaluation of ID impact.[9]

1.08 *Scale:* Problems of attribution are compounded by problems of aggregration—that is, summing up the impact of numerous projects on the overall quality of a country's public management architecture. One approach is simply to assume the whole is equal to the sum of the parts: Simple non-weighted average of ID ratings from individual Bank operations could serve as country-level metrics. However, this approach is inadequate because it assumes that all institutions serve the same purpose or register the same impact on public management systems.

---

[4] OED defines institutional development (ID) impact as "the extent to which a project has improved an agency's or a country's ability to make effective use of its human and financial resources." OED, 1997.
[5] North, 1990.
[6] Riker, 1982
[7] Shepsle, 1998.
[8] Goodin, 1996; Stevens, 1993; Riker, 1982.
[9] World Bank, 2000.

To the contrary, assigning relative weights to project-level ID ratings necessarily requires a series of hypotheses about how discrete project-level interventions at one level of the state may affect institutions at another level. One hypothesis, for instance, is that projects that fail to improve the predictability of resource flows actually hamper the ability of line managers to contract out service delivery functions to private providers. Another may be that short-term funding mechanisms designed to rapidly disburse development assistance off-budget can fragment planning processes, weaken financial accountability, and undermine allocative efficiency. Evaluators then will need to develop metrics that reflect both the intended and unintended effects of ID efforts on public management systems.

1.09 Implicit in this type of *systems-based* approach is an acknowledgment that ID is neither linear in trajectory nor cumulative in effect. Successful project-level interventions do not necessarily add up to improvements in a country's overall institutional endowment. Rather, they are constrained by a range of very real trade-offs, such as those between achieving short-term frontline performance gains and longer-term sustainable improvements in public management, between increasing the predictability of inputs and allowing greater flexibility in their use, etc. The task for evaluators then is to make these trade-offs explicit so that the Bank and its clients are able to make well-informed and strategic choices about the redesign of public institutions.

Objectives and Organization of the Paper

1.10 Problems in defining and measuring institutions, attributing institutional change to various exogenous and endogenous factors, and aggregating the effects of multiple interventions on systems are not new. They have troubled scholars and practitioners for years and are, in some sense, intrinsic to evaluation. With that in mind, the objective of this piece is modest: to offer some basic analytical tools that evaluators can use to assess—in a standardized manner—the impact of Bank policy advice and lending operations on a key component of a country's institutional endowment, namely, its system of public management.

1.11 To this end, **Section 2** delineates that subset of "public institutions" with which CAEs should be most concerned. It draws on the recent literature in the economics of collective choice and organizational design to provide a framework for evaluating institutional quality, explaining performance, and understanding the process of change. **Section 3** uses this framework to assess the treatment of public management issues in past country evaluations by OED. Based on lessons from this self-assessment, **Section 4** and relevant annexes suggest an approach for evaluating the country-level impact of Bank support for public management, including methodologies for "institutional mapping", evaluating relevance and responsiveness, rating efficacy, and attributing the outcomes of ID efforts. Finally, **Section 5** offers a brief conclusion with recommendations for mainstreaming the proposed methodology into OED's existing evaluation architecture.

2. Thinking about Public Management

2.01 This section sets forth a definition of institutions around which there is growing consensus. It also identifies those elements of public management that are of particular interest to CAEs. It then synthesizes recent thinking on "good practice" approaches to designing these public institutions as well as key endogenous and exogenous factors that typically contribute to institutional change.

Institutions and Organizations

2.02 That institutions matter for development is well-established. The contemporary literature defines "institutions" as the formal and informal rules that govern production and exchange and thereby endow these interactions (particularly repeated ones) with the stability necessary for human beings to function with "at least a modicum of rationality."[10] A society's institutional endowment also includes "organizations"—individuals and groups organized around a more or less common purpose—which emerge in response to the prevailing incentive environment created by the "rules of the game."[11] Organizations (and entrepreneurs within them) exercise choice in following these rules and, at times, seek to change the rules to further their own interests. Effective organizations are those that employ physical, financial, and human assets in ways that further their mission and objectives within a given institutional environment.

2.03 Of particular interest to development practitioners is *the process* or *the mechanics of institutional change*. The nascent literature on this topic suggests that institutional structures are continually subject to optimizing and opportunistic behavior, coalition formation and capture, time inconsistency, and other rapidly changing preferences. Even as they face these myriad pressures to change, institutions are weighed down by the inertia of existing equilibria (some of which may be stable, yet dysfunctional). From this vantage point, *structural reforms* can be seen as deliberate attempts to dislodge institutions from sub-optimal equilibria and re-form them within governance structures that provide stronger incentives to perform.[12] By the same token, *capacity building* efforts are attempts to supply organizations with skills and equipment (or other inputs) to operate more efficiently within a given institutional context. It is also worth noting that attempts to intentionally redesign institutions or build organizational capacity may also set in motion unintended dynamics, some of which could adversely affect the very structures that are the focus of reforms.

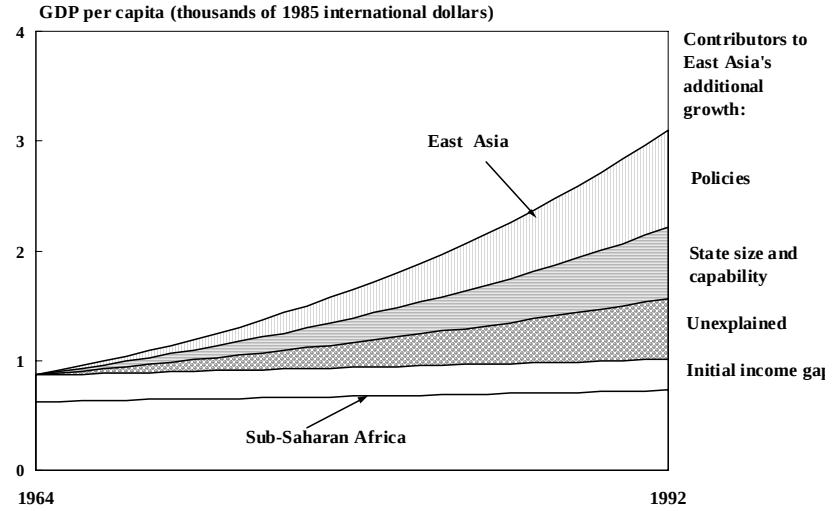
Which Institutions Matter and Why

2.04 Virtually every policy intervention—whether it be exchange rate liberalization or outsourcing of works—has some effect on a country's institutional stock. However, measuring the full range of institutions that directly and indirectly affect growth and development is impractical, if not impossible. Given finite resources and time, evaluators are likely to assess ID impact in areas identified as priorities for a specific CAE (for example, the financial sector, the regulatory framework for the private sector, or social service delivery). That said, it is worth noting that performance in health, education, or infrastructure, as well as cross-cutting concerns such as fiscal performance, all depend on the quality of public management. Therefore, systems and structures that help transform *inputs*—garnered through public resources—into *outputs* and *outcomes* should rank high among the "institutions that matter most" for CAEs.

[10] Simon, 1983.
[11] North, 1990.
[12] Aoki, 1998; Dasgupta, 1999.

2.05 The significance of public management in shaping countries' development prospects is further reinforced by a growing body of empirical work. According to the *World Development Report 1997,* "capability"—or the ability of the state to undertake collection actions at the lowest cost to society—explained a significant portion of the divergence in economic *outcomes* between Sub-Saharan African and East Asian countries over a 30-year period (Figure 2.1).[13] OED's *1997 Annual Review of Development Effectiveness* also concluded that the performance of Bank-financed projects improved with the capability of government institutions.[14]

Figure 2.1: Public Management Explains Part of Divergence in Economic Performance



Source: **World Bank, 1997 World Development Report**

2.06 In addition to the fact that "institutions matter," empirical analysis points to specific types of institutional arrangements that are associated with "good" public management outcomes. For instance, evidence from 22 Latin American and Caribbean countries indicates that the governments that employed hierarchical and transparent processes for budgeting were more likely to avoid large fiscal deficits than countries with collegial and non-transparent budgeting processes.[15] Alternatively, devolution has been associated with demonstrable reallocation of resources to priorities that reflected local preferences, which also happen to be higher return activities.[16] Similarly, outsourcing and contracting arrangements have been shown to increase output efficiency as well as client satisfaction in infrastructure and some social sectors.[17] Taken together, this empirical work has helped identify specific types of arrangements that best facilitate the transformation of inputs into outcomes—a process that lies at the heart of public management.

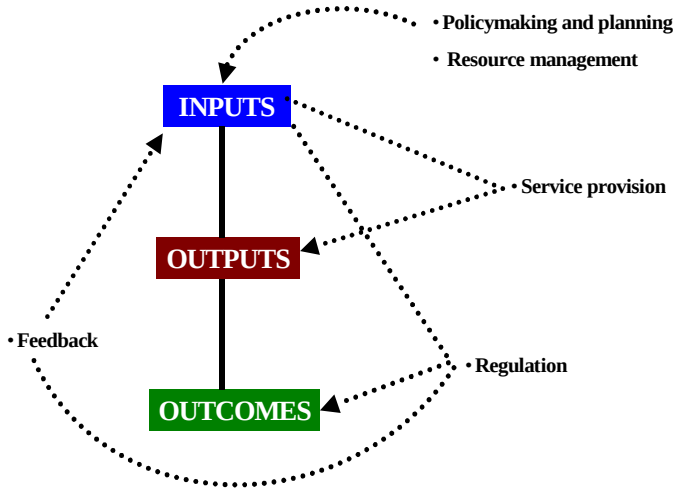
The Transformation Process in Public Management

2.07 Public management systems are principally concerned with the transformation of inputs into outputs and outcomes (Figure 2.2).[18] The incentive framework governing transformation comprises a range of institutions such as formal rules, standard operating procedures, legal covenants, as well as norms, customs, and unwritten codes of conduct. Myriad organizations at

[13] World Bank, 1997.
[14] Operations Evaluation Department, 1997.
[15] World Bank, 1997; Alesina et al., 1996; Campos and Pradhan, 1996.
[16] Faguet, 1998.
[17] Shah, 1998; Bennett et al., 1997; Hanushek, 1994.
[18] Adapted from seminar by Andrew Lawson, DFID Governance Retreat, Redwood, UK, September 1999.

different levels work to support this transformation process by formulating policy, allocating resources, undertaking service delivery operations, and enforcing the rules that govern such activities.[19] Together, policymaking and resource management processes (for example, budgeting and staff recruitment) help set aggregate levels for input use as well as allocation of financial and human resources between competing priorities (for example, defense and education). Institutional arrangements governing delivery shape the productivity of operational activities undertaken by line agencies and other organizations that convert inputs into monitorable outputs. Finally, various regulatory mechanisms help safeguard transformation from arbitrariness (such as corruption at the point of delivery or patronage in recruitment), monitor output performance, and evaluate links between outputs and outcomes. Lessons from evaluation provide important feedback that actors at various points in the transformation process can internalize and use to improve processes, systems, and structures. The performance of organizations that carry out these functions (policymaking, resource management, delivery, and regulation) profoundly shapes the effectiveness and efficiency of transformation. The literature on organizational performance is too vast to summarize here. For the purposes of this paper, it will suffice to say that organizations—whether they are involved in regulation, budgeting, or service delivery—tend to perform at higher levels when supported by appropriate *structural* and *capacity* conditions.[20]

Figure 2.2: Transforming Inputs into Outcomes



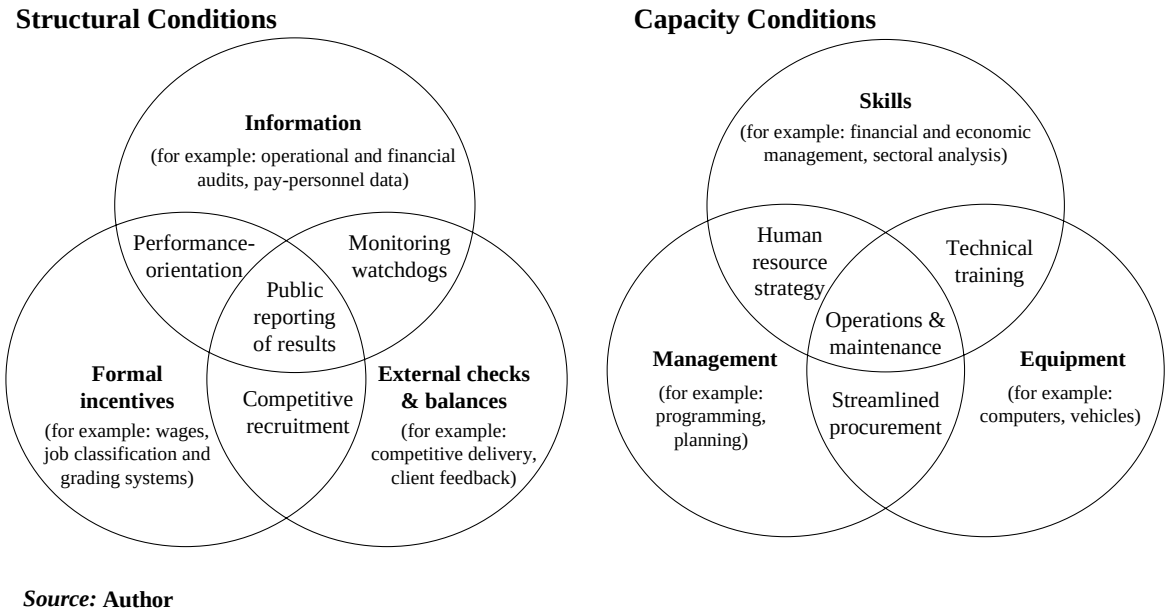
Source: **Adapted from Lawson, 1999.**

2.08 *Structural Conditions.* Governance structures shape the incentive environment within which organizations, and individuals, undertake critical activities related to public management such as service delivery or budgeting. According to the organizational design literature as well as the Bank's own experience in client countries, three structural elements—formal incentives, external checks and balances, and the availability of timely information—are considered to be conducive to improved performance (Figure 2.3). *Formal incentives* typically comprise wage remuneration competitive with the private sector comparators, as well as job classification and grading systems consistent with internal job relativities.[21] They also include formal rules that reward or sanction various practices or types of behavior. When they are enforced, rules and restraints can also help promote predictability and limit arbitrariness in critical activities such as budget formulation and execution or personnel management. In addition, there is growing acknowledgment that *external checks and balances* such as participation or competition can bring client feedback and related pressures to bear on government agencies, limit discretion of bureaucrats, and provide for the credible threat of exit or citizens opting for alternative suppliers—all of which encourage public servants to perform better. Finally, disseminating *information* on performance—such as the cost efficiency of health facilities or client satisfaction data—is likely to elicit a demand-side accountability response from clients.[22]

[19] Girishankar and Manning, 1999.
[20] Africa Capacity Building Group, 1999; also see World Bank, 1997.
[21] Lindauer and Nunberg, 1994; Lienert and Modi, 1997.
[22] Williamson , 1985; World Bank, 1998.

2.09 Each of these mechanisms in and of itself promotes performance. When combined, they can be complementary and mutually reinforcing. For instance, governments can open civil service recruitment to competitive pressures not only by providing wages that are comparable to the private sector, but also by encouraging top cadres to migrate from tenured positions to renewable fixed-term contracts. At the same time, criteria for renewal of contracts can be defined in terms of good stewardship of inputs, specifically requiring civil servants to fulfill the hierarchical reporting requirements related to financial management. Another complementary mechanism to promote performance could be to increase significantly the availability of information on how effectively public resources are used by strengthening independent watchdogs, supreme audit institutions, and related mechanisms of parliamentary oversight.

Figure 2.3: Structural and Capacity Conditions for Institutional Performance



Source: **Author**

2.10 *Capacity Conditions.* Public sector organizations operating within appropriate governance structures also need a minimum level of know-how, skill, and equipment—"software and hardware"—to fulfill their objectives. Weak capacity is an acute problem in many developing countries. In many low-income countries, civil servants are not equipped with even basic complementary inputs such as stationery, furniture, computers, and other specialized technology to carry out their day-to-day tasks.[23] Such instances are reflective of environments characterized by pervasive aggregate scarcity. However, their persistence over years further highlights the urgent need for innovative approaches to injecting knowledge and assets into weak-capacity environments. Past experience suggests that donor financing of training or equipment to fill short-term capacity gaps is often inadequate and unsustainable in the absence of effective management of resource flows. Rather, success in such public sector organizations depends heavily on the development of managerial talent, not only to administer projects and programs, but to continually navigate problems of budget uncertainty and pervasive scarcity.

2.11 Public management systems will more effectively transform inputs into outputs and outcomes when the design of "good practice" governance structures (such as those described above) is properly sequenced with targeted efforts to build capacity. In recent years, many "capacity building" activities have come under disrepute in part because lopsided interventions

[23] Girishankar, 1997.

actually injected human and physical resources into structurally deficient settings, with little impact. A 1999 OED review found that the Bank often supported training activities and the procurement of equipment even when civil services lacked incentives (for example, competitive wages and performance-monitoring systems) to channel and utilize capacity efficiently. The net result was to encourage "the migration of talent," as higher-level staff, who received training, leveraged new skills to secure employment in the private and donor sectors.[24] Conversely, this implies that efforts that first restructure institutional arrangements, and then supply skills and equipment, are most likely to drive organizational performance.

2.12 Public sector organizations need to meet different structural and capacity conditions (for example, participatory processes are more appropriate than hierarchical rules) in order perform well. These conditions depend on the type of function or activity that organizations are required to undertake. For example, units involved in budget execution may respond favorably to hierarchical structures that do not allow for much discretion or capture by special interests. Service delivery organizations, however, may be more efficient and accountable under more open and participatory structures that allow for greater client ownership and involvement. What follows is a more thorough description of the core elements or functions of public management and the structural and capacity conditions that would best support them.[25]

Elements of Efficient Transformation

2.13 Experience and the growing body of empirical work in developed and developing countries suggest that the transformation process is most efficient and effective (Figure 2.2) when characterized by (i) *credible* policymaking and planning, (ii) *adequate* and *predictable* resource flows, (iii) *flexible* delivery arrangements, and (iv) *enforceable* regulation. These four attributes of effective public management systems are discussed below, along with the structural and capacity factors required to promote them.

Credible Policymaking and Planning

2.14 From its inception, policy-based lending by the Bank has sought to promote technically sound macroeconomic and sectoral policies that would result in rapid economic growth and poverty reduction. The assumption behind this approach to assistance was that development was ultimately a matter of good policies, which could be developed under the stewardship of well-intentioned political leaders and benign, highly trained technocrats.[26] The adjustment experience, however, suggests that bright technocrats, armed with big ideas, do not by themselves guarantee the quality, consistency, and stability of economic policy. Rather, policies are more likely to be sustainable and irreversible when they are credible in the eyes of cabinet-level actors, civil servants in central and line agencies, and broad constituencies outside government. More important, credible commitments to policy proposals depend as much on the *process* as they do on analytical and technical content. The institutional arrangements that shape *how* policies are made therefore profoundly influence their content and credibility.[27]

2.15 To be effective, assistance for policy reforms should include support for institutional arrangements and processes that ensure formulation of (i) macroeconomic policies that maximize the "size of the [economic] pie," and (ii) sectoral policies that efficiently "divide the pie."[28] Putting aside questions of substantive policy content, experience suggests that both monetary

[24] OED, 1999.
[25] Preker et al., 1999.
[26] PREM, 1999.
[27] Biddle and Milor, 1999.
[28] Sectoral policymaking hinges on the credibility of monetary policy and fiscal aggregates. Manning et al., 1999.

policy as well as the setting of the aggregate fiscal framework are best carried out by organizations that are *insulated* from the exigencies of everyday politics and, therefore, the risks of capture. If macroeconomic policies (both monetary and aggregate fiscal) were formulated using participatory processes, they could be easily captured by actors with high discount rates, such as political incumbents, the elderly, and other groups who are more willing to forgo long-run stability in return for short-term gains. Capture by such special interests, which typically lack the technical expertise to management monetary or fiscal policy, could severely distort long-run growth prospects. There is broad consensus among practitioners and researchers that macroeconomic policy is best formulated in an insulated manner, specifically by "delegating monetary policy to an independent Central Bank, [or] establishing technocratic enclaves within the Ministry of Finance, fixed policy rules, and binding international agreements."[29] Such arrangements would ensure that highly technical problems of the macroeconomy are managed by qualified experts in a de-politicized setting.

2.16 Once governments provide an institutional framework in which policies that expand the "size of the pie" are formulated, they need to address the ancillary question of sectoral policymaking or "dividing the pie." There are three reasons why inclusive and participatory institutions—as opposed to insular ones—are prescribed for sector policymaking, or the strategic prioritization of how public resources are spent. First, "dividing the pie" efficiently requires that sectoral policies give due consideration to inter-sectoral trade-offs as well as intra-sectoral ones (such as those between investment and recurrent expenditures). However, information asymmetries associated with strategic prioritization imply that a highly technocratic, insulated approach to sectoral policymaking would be both infeasible and ineffective. For this reason, the participation of beneficiaries, local and provincial governments, external groups, as well as bureaucrats is necessary to "divide the pie" efficiently. In addition to allowing various economic actors, particularly the poor, to express their preferences, more consultative policymaking processes allow for open debate over costed policy alternatives. Second, inclusive and open processes also increase the likelihood that politicians (particularly cabinet-level actors), once publicly committed to a set of sectoral policies, will not renege. If they do renege, politicians could face pressure from organized groups with longer-term interests in specific policies and priorities.[30] Finally, involvement in planning processes enables local citizens—especially members of excluded groups—to cultivate certain expectations, even a sense of entitlement, about the content and performance of government policies. In the process, local communities are mobilized to act as regulators and evaluators during implementation.[31]

2.17 Insular and inclusive arrangements—each in their own way—satisfy the respective structural conditions for sound macroeconomic and sectoral policymaking. The capacity requirements for making "good policies" are also considerable. These include the development of skills to undertake high-level planning and economic management functions as well as design and implement sector-specific programs. Investment in the statistical base, relevant information systems, and computer hardware for monitoring various development outcomes are also critical. Table 2.1 summarizes the structural and capacity requirements for sound policymaking; these should form the basis for evaluating development assistance in areas such as macroeconomic and public expenditure management as well as policy reform within particular sectors.

Adequate and Predictable Resource Flows

2.18 Robust arrangements for policy formulation and strategic prioritization would be for naught without some assurance of adequate and predictable resource flows during budget

[29] Biddle and Milor, 1999.
[30] Shepsle, 1998.
[31] Perry et. al., 1999.

execution. Budgetary uncertainty can result when agreements between cabinet-level actors are not mutually binding.[32] It also derives from over-optimistic revenue projections within the aggregate fiscal framework, as well as weak commitment control systems, poorly designed releases from the cash budget, and a lack of transparency during budget execution.

2.19 As noted earlier, hierarchical, transparent rules for budget execution are more likely than non-transparent collegial arrangements to prevent problems of arrears, overruns, and, therefore, large fiscal deficits. There are several ways in which hierarchy and transparency can be used to ensure predictable and adequate resource flows downstream of the budget cycle. One approach, used in OECD countries such as Australia, is to institutionalize a system of forward estimates or cost estimates for policies carried forward into the medium term.[33] Every year, the government publishes both these forward estimates as well as proposed allocations to make transparent any divergence. Past financial and operational performance are factored into the elaboration of future estimates for specific policies and programs.[34] This mechanism provides an incentive for line ministries—heavily involved in a transparent process budget process—to ensure that programs are delivered within the available resource envelope. Such "lock-in" mechanisms also deter cabinet-level actors from making ad hoc re-allocations in the middle of the budget cycle. Taken together, efforts to ensure transparent formulation processes and execute the budget in line with forward estimates all help make the budget a credible signal of government policy.

2.20 Several developing countries are also attempting to replace incrementalist budgeting with medium-term expenditure frameworks (MTEFs) that operate on a three-year rolling cycle. Recent experience with these MTEFs suggests that execution is the fulcrum on which the credibility of the budget process turns. In several countries, sector ministries have little incentive to participate in elaborated medium-term planning exercises coordinated by the Ministry of Finance unless they can be certain that resources will be made available as planned. Effective execution requires a robust financial management infrastructure comprising *inter alia* modern, organic finance legislation, Treasury instructions, other financial management procedures, and computerized financial management information systems to ensure up-to-date financial accounting and reconciliation of below-the-line items and bank accounts. Equally important are incentives for the head civil servant in each line ministry to carry out his or her responsibilities as controlling officer. Accordingly, the development of a skills base in public sector accounting and auditing, as well as strengthening of supreme audit institutions, remain priorities.[35]

2.21 In addition to adequate and predictable budgetary resources, well-functioning governments require a capable and motivated staff. The legal and governance arrangements for hiring and firing of civil servants, the design of job classification and grading systems, as well as codes of conduct shape the backbone of government personnel systems. Statutory bodies such as public service commissions and other appeals bodies also serve to preserve the political neutrality and professionalism of various cadres of public servants. Over the past two decades, the Bank has been intensively involved in efforts to strengthen civil service systems in developing and transition countries. Two sets of civil service problems were found to be common across countries: (i) overstaffing of civil service cadres (particularly at the lower levels), driven by populism and patronage, and (ii) migration of managerial and technical staff out of the civil service because of poor real wages.[36] The resulting vacancies in key senior and technical posts tend to persist until internal wage relativities (between higher and lower grades) are decompressed and also wages for important cadres are brought in line with private sector

[32] Manning et al., 1999.
[33] Campos and Pradhan, 1996.
[34] Ibid., 1996.
[35] Stapenhurst and Dye, 1999.
[36] Girishankar, 1997; OED, 1999.

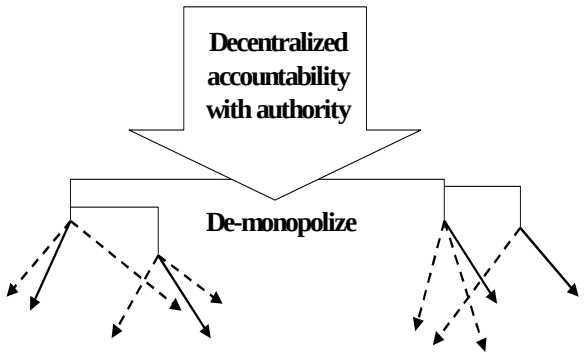
comparators. In addition to competitive wages, governments need to ensure that staff-in-post are equipped with the necessary inputs such stationery, vehicles, and information technology to carry out their tasks. Another important aspect of Operations and Maintenance (O&M) is provision in unit budgets for training and career development of civil service cadres. Effective establishment control and payroll management also require accurate information on staff size (para 2.31).

Flexible Delivery

2.22 Credible policies, predictable budgets, and motivated staff are the mainstay of good public management, but they alone do not guarantee efficient service delivery. The design of delivery systems is critical to improved access to services such as healthcare and education, utilization rates, operational efficiency of facilities such as clinics and schools, and finally, client satisfaction. Several developed and developing countries have attempted to develop more innovative approaches to financing, delivering, and monitoring infrastructural and social services. Fiscal crises and popular frustration with centralized, monopolistic systems provided the impetus for these far-reaching experiments in service delivery. Through the 1980s, practitioners in the OECD responded by applying lessons from the corporate restructuring experience—for example, unbundling activities such as finance and production, opening them to competition, and decentralizing decisions to frontline managers—in order to deliver public services more efficiently, particularly those at the local level.[37] Others attempted to ride the wave of decentralization in the 1990s that would sweep across Latin America, South Asia, and Africa. Decentralization was seen as an opportunity to improve allocative and technical efficiency by shifting fiscal and administrative responsibilities to the lowest possible levels of government, in accordance with the subsidiarity principle.[38] Suffice it to say, by the mid- to late1990s, the public manager's "toolkit" comprised a variety of new contracting arrangements—from corporatization to devolution—to improve cost efficiency and client orientation.

2.23 The key behind innovations in service delivery was that line managers were empowered with greater *flexibility* to produce outputs efficiently. Specifically, flexibility entailed two structural reforms—first, the decentralization of *accountability* to lower levels of government along with the *authority* to control fiscal and administrative inputs, and second, the de-monopolization of output production by employing a richer combination of hierarchical rules, competition and voice (Figure 2.4).

Figure 2.4: Two Dimensions of Flexibility



Source: **Author**

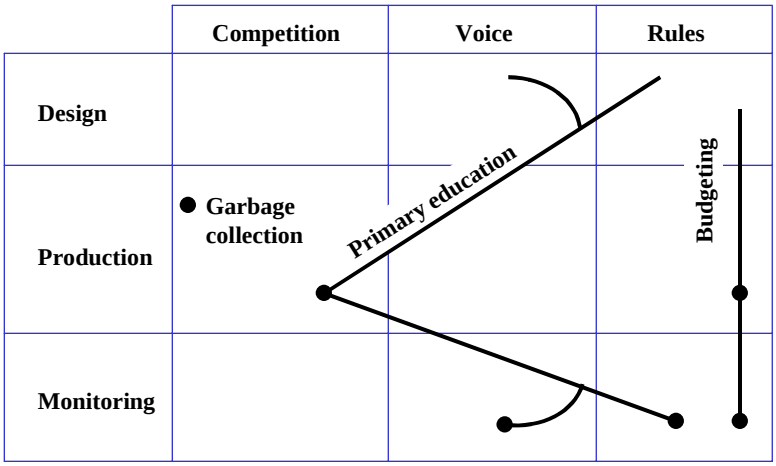
2.24 *Decentralization.* Decentralization has found many advocates in recent years. Its purported benefits include increased efficiency in the allocation of financial resources, improved access to the rural poor, and greater opportunity to exploit economies of scope. In contrast to centralized delivery where accountability for results tends to be diffused, decentralized systems ensure a greater degree of proximity between citizens and public officials; as a result, the lines of accountability downward to citizens are much clearer. International experience indicates that such arrangements are associated with improved coverage and utilization, as well as resource allocation in line with local preferences.[39]

[37] Sabel, 1995; Bennett et al., 1996.
[38] World Bank, 1999.
[39] Faguet, 1999.

2.25 While the body of empirical work in this areas is still modest, preliminary evidence suggests that decentralization alone does not generate economic competition or induce people to "vote with their feet," as theory suggests. Rather, decentralization by itself can reproduce the same monopolistic, and vertically integrated arrangements that characterized centralized systems, with associated problems of poor incentives, inefficiency, and arbitrariness. For local authorities, the prospects of taking on the responsibility for service delivery *without* the authority to de-monopolize that function—including the freedom to experiment with new techniques for output production—are neither feasible nor desirable.[40] However, the benefits of decentralized delivery are enhanced when vertically integrated activities such as funding, delivery, and regulation are unbundled at the local level and then subjected to competition or other pressures that promote efficiency and client orientation.

2.26 *De-Monopolization.* The prescribed method of de-monopolizing a service delivery activity—that is, unbundling the activity and subjecting it to various forms of competition or competition surrogates—depends *inter alia* on the production characteristics of the good or service.[41] Certain outputs such as garbage collection can be easily observed, and their asset specificity is low enough to allow the entrance of alternative suppliers. For these contestable and easily monitored outputs, competition in a spot market is the appropriate mechanism for delivery. Other activities such as budgeting are hard to measure and difficult to subject to competition due to high asset specificity. As depicted in Figure 2.5, such activities should necessarily be carried out within the core public sector using transparent and hierarchical rules. Activities for which pure hierarchy is the appropriate mechanism are rare.

Figure 2.5: Mapping Options for De-monopolization



Source: **Adapted from Girishankar, 1999.**

2.27 Most outputs such as those in social sectors are relatively contestable, but difficult to measure. In order to improve their ability to monitor output performance, some governments have started to rely on the comparative strengths of technical experts as well as beneficiaries in observing output performance in the social sectors. For instance, parents can observe whether micro-level interventions made by a school actually improve their child's interest in education, day-to-day performance, and progress over time. Technocrats, on the other hand, are better able to ensure that both content and testing procedures are modern and in line with advances in knowledge. What this suggests is that education requires a design phase—involving both the participation of parents and teachers as well as technocratic inputs from experts—to define the content and performance standards for "public education." Once standards are defined, education can be subjected to competition among private, non-profit, and public providers. One option is a voucher or charter system in which "the money follows the child" to the school of his or her choice. Participation in this system, however, requires that the school, regardless of its legal

[40] Burki et al., 1999, for discussion of health services decentralization in Latin America.
[41] Girishankar, 1999.

status, meet the publicly mandated standards prescribed in the design phase.[42] Once delivered, the service is monitored and evaluated by multiple stakeholders according to their particular comparative advantage. Figure 2.5 illustrates how flexible service delivery arrangements allow for a rich mosaic of actors and institutional arrangements in the design, production, and monitoring of social sector outputs.

2.28 Since the late 1980s, the Bank has made some progress in supporting de-monopolization, particularly in its assistance to infrastructural sectors such as telecommunications, power, and roads. It has yet to systematically support flexibility-enhancing reforms—combining decentralized control of inputs and de-monopolized productions of outputs—in the social sectors. For a variety of reasons, including lack of commitment to institutional reforms, the Bank's portfolio includes several operations designed to improve the performance in monopolistic, centralized service delivery agencies through capacity building efforts.[43]

2.29 Countries that have undertaken structural reforms to enhance flexibility invariably face pressures to build capacity in financial management, procurement, information technology, regulation, and general management. For instance, civil servants will need increasingly to be familiar with competitive procurement of consultancy as well as goods and services contracts. They will need to develop techniques and methodologies for benchmarking performance, and familiarize themselves with procedures for regulating multiple providers. As civic and private actors attempt to contribute as producers of public services, they too will need to develop basic skills in finance and accounting, personnel management, and bidding. Regardless of the specific capacity needs of public and private actors, it will be important for the Bank operations to employ more demand-driven and cost-efficient approaches to training. For example, consortia of non-governmental organizations or local governments can set up a consulting body (such as Service Public 2000, established by municipalities in France) that trains and consults with public and private organizations for a fee.[44] Twinning arrangements can be established between institutes in developing countries and those in the OECD in order to provide relatively inexpensive training and career development opportunities.

Enforceable Regulation

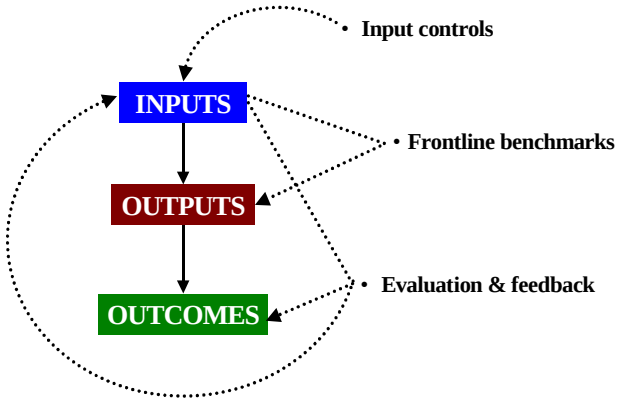
2.30 The various elements of transformation—including setting fiscal aggregates, allocating across strategic priorities, producing outputs efficiently, and generating desirable outcomes—do not proceed automatically. Concerted and deliberate efforts are required to ensure that "collective actions" such as service delivery or budgeting are undertaken at minimum cost to society, or with minimal risk of arbitrariness, inefficiency, and ineffectiveness.[45] Safeguarding the transformation process from such risks implies the need for a regulatory regime the covers the management of inputs, the production of outputs, and the generation of outcomes. Regulation should also provide incentives for public actors to learn, adapt, and innovate activities in each phase. Described below are the structural and capacity aspects of regulatory arrangements recommended for public management systems (Figure 2.6).

2.31 *Checking Arbitrariness in the Management of Inputs.* When public actors enjoy too much discretion in the use of inputs without countervailing checks and balances, they are prone to various forms of opportunism and arbitrariness. Moral hazard problems in resource management,

[42] Private and civic institutions are increasingly involved in activities (financing, delivery, and regulation) that contribute directly to the public interest. Under such circumstances, organizations can be deemed "public"—regardless of their proprietary or legal status—if they meet common standards for content and performance. Girishankar, 2000.
[43] Ibid.
[44] World Bank, 1999.
[45] Chhibber, 1998.

if unchecked, can not only threaten government's fiduciary interests but also undermine frontline performance and development outcomes. Accounting standards, robust commitment control procedures, internal and external auditing, and some form of legislative oversight are examples of standard *input controls* that limit discretion and ensure due process in the resource allocation and management. Donors and other members of aid consortia have increasingly acknowledged the importance of external auditing by semi-autonomous bodies such as supreme audit institutions (SAIs) in deterring opportunistic behavior by the executive. This type of check and balance, which usually includes the dissemination of information on financial performance, can be adapted to different legal, political, and administrative settings. Even in decentralized settings such as Uganda, tracking of primary education expenditures and dissemination of the results has generated a demand-side response from communities, which have started to apply pressure regularly on school principals, bureaucrats, and frontline workers to account for the public resources they receive. At the systemic level, efforts to improve parliamentary oversight in several anglophone countries have typically involved strengthening the independence of Offices of the Auditor General as well as the Public Accounts Committees. The capacity of audit bodies and citizen groups to enforce input controls is an ongoing concern. Formal, professional training of auditors and functional education for citizens are two capacity conditions that should be met for accountability mechanisms to work effectively. Similarly, investments in human resource information systems such as those that link pay and personnel should facilitate monitoring of payroll, personnel systems, and the financial implications of personnel decisions.[46]

Figure 2.6: Safeguarding and Improving Transformation Through Regulation



Source: **Author**

2.32 *Continuous improvement through benchmarking outputs.* As discussed in para. 2.23, prospects for output efficiency depend in large part on the flexibility of delivery arrangements. However, flexible systems—or those that have decentralized and then de-monopolized the production of service delivery outputs—come with their own set of risks, such as shirking by private providers, various forms of corruption, and information asymmetries that make monitoring difficult. The regulatory dilemma for public management systems is how to promote flexibility while simultaneously guarding against contractual breakdowns and the resultant delivery failures that could be regressive in impact. One approach, international experience suggests, is first to *set minimum standards* or "floors" below which performance will not be permitted to fall, and then to *benchmark* output performance (and publish the same). *Benchmarking* outputs (Table 2.1) fosters inter-jurisdictional competition, attracts potential private investors, and raises the costs of "harboring" poor performers. In flexible systems, service providers would have the freedom to experiment with the *means* of delivery and test new techniques, processes, or technologies in order to meet or surpass output benchmarks or standards.[47] Every time they are met, benchmarks are ratcheted upwards so that delivery systems are subjected to ever rising performance floors but not performance ceilings. Even as governments invest in the skills and know-how to establish standards and benchmarks, they

[46] Nunberg,1995.
[47] Sabel, 1995.

should seek to match the design and sophistication of these systems to local conditions, particularly the availability of economic data.

2.33 An important caveat is that the publication of output performance (vis-à-vis benchmarks) is necessary, but not sufficient, to ensure accountability for results. What is critical is the ability of published information on output performance to elicit a demand-side response from beneficiaries. Beneficiaries are more likely to demand accountability when service providers have failed to meet their expectations or their perceived entitlement. One way to raise expectations (and increase demand for information) is to ensure broad-based participation of technocrats, providers, and beneficiaries in setting benchmarks or content and performance standards. Participation by beneficiaries ensures that standards are demand-driven, more realistic given local constraints, and more easily monitored.

2.34 *Understanding results.* Policies and programs are premised on the notion that specific outputs in a given sector contribute to outcomes, considered desirable for growth, development, and wellbeing. For instance, certain recommended pedagogical techniques are believed to result in higher levels of literacy. An educated workforce in turn is able to contribute to higher levels of productivity and broad-based growth. Such hypotheses regarding the economic benefits education have been empirically verified using longitudinal, cross-country data. In other sectors, hypotheses regarding the impact of certain outputs on outcomes may not be well established. In such cases, government policy may actually be financing outputs—the "wrong" outputs—that have only a marginal, or even negative, effect on desired results. In still other cases, even when the "right" outputs are produced, results are undermined either by technical inefficiencies that are *endogenous* to the delivery arrangements (for example, high administrative costs associated with publicly managed hospitals) or by unforeseen *exogenous* factors. Monitoring and evaluation (M&E) systems enable policymakers to determine whether the achievement of desirable outcomes requires the provision of different outputs, more efficient implementation of existing programs and policies, or new types of interventions to address previously unforeseen factors.

2.35 The most basic requirement for rigorous evaluation is data on inputs, outputs, and sectoral outcomes. Since most Bank operations are implemented in information-poor and capacity-weak settings, methodologies devised for public sector M&E systems should be simple. When relevant and feasible, multidisciplinary evaluation frameworks should be better able to attribute development outcomes to specific policies and programs or exogenous factors. Finally, the capabilities of public sector evaluators and auditors—including their ability to play a more prominent role during budget formulation (para. 2.38)—are critical if public managers are to effectively internalize the lessons from past development efforts. Finally, investments in building and updating the statistical base as well as developing a career track for professional evaluators are some capacity-side measures that should be considered.

2.36 Effective M&E is not simply a matter of rigorous analysis. The structure and governance of evaluation play critical roles in bringing evaluation results to bear on government operations and the overall direction of policy. Three structural variables—(i) the independence of (and incentives within) evaluative and audit bodies, (ii) the inclusiveness and openness of evaluative processes, and (iii) links between evaluation and key allocative or technical decisions—should be considered in evaluating the adequacy of M&E systems.

2.37 In some countries such as India, the independence of SAIs is constitutionally mandated and operationalized through regular reporting to a parliamentary Public Accounts Committee. Even formal independence, however, can be undermined by conflicting reporting and funding relationships. For instance, bodies such as SAIs can be weakened if they receive the bulk of funding from the very institutions that they are to audit (for example, the Ministry of Finance).

Independence is also conditioned by the auditing institution's legal authority to access any and all information from operational units of government. In some countries, evaluators have limited access to classified information, which limits the types of audits that can be credibly undertaken.[48] Finally, it is important that evaluative and audit institutions reward their staff with competitive remuneration packages and insulate them from political interference so that quality and impartiality are assured.

2.38 As note above, structural independence of evaluation bodies furthers the accountability objectives of M&E systems. More inclusive arrangements, however, are required to further the learning objectives of M&E; these typically encourage interactive participation by evaluators, the evaluated, and the beneficiaries in the review of policies and programs. Prospects for learning and innovation are also enhanced when self-evaluation is routinized in line agencies themselves.[49] Finally, the full impact of M&E is realized when it is explicitly linked to the budget process, and specifically, to decisions about resource allocation in the future and the redesign of existing government operations. In countries that have placed high priority on monitoring outcomes, such as Australia, consideration of evaluation results is central to cabinet-level decision making, policy

Table 2.1: Elements and Indicators of Efficient Transformation

Elements	Structural Conditions	Capacity Conditions	Institutional Quality Indicators	Front-line Performance Indicators
CREDIBLE POLICYMAKING	• Delegation to technocratic enclaves for macroeconomic policymaking • Participation, open debate over costed alternatives for sectoral policymaking	• Skills in economic and policy analysis • Robust database on key economic variables	• Volatility • Inconsistency • Incoherence	• ACCESS • UTILIZATION • TIMELINESS • SATISFACTION • COST EFFICIENCY
ADEQUATE AND PREDICTABLE RESOURCES	• Hierarchical and transparent budget processes • Competitive terms of services (wages, grading and evaluation systems) • Binding restraints on cabinet decision-makers • Wage and non-wage incentives	• Skills in budgeting, financial management, procurement • Well-trained staff in post for budgeting, personnel management • Information on pay-personnel, financial management	• Budget variance • Composition of recurrent and capital expenditures • Degree of leakage • Vacancy rates in key posts	
FLEXIBLE DELIVERY	• Decentralize authority to control inputs with accountability for results • De-monopolization to introduce voice and choice	• Capacity in local government, among private providers • Statistical base, particularly of clients • Hardware, infrastructure	• Unfunded mandates • Standards over outputs, not inputs or processes • Number of alternative providers per jurisdiction	
ENFORCEABLE REGULATION	• Arms-length, independence of auditors and evaluators • Inclusive and open evaluative processes • Tight links between evaluation and resource allocation	• Skills in auditing, survey design, and evaluation • Information network for 'live databases' from sectors • Adequate capacity and resources for independent bodies	• Incidence of corruption and resource leakage • Frequency in changes in the formal rules • Learning by doing • Changes in policies and program design due to M&E findings	

Source: **Author, with adaptations from Manning and Knack, 1999.**

[48] Stapenhurst and Dye, 1998.
[49] Bhatnagar and Williams, 1992.

formulation, and budget preparation.[50] Similar efforts to improve the integration of M&E into public expenditure management are underway in countries such as Zimbabwe, Guinea, and Uganda.[51]

2.39 Table 2.1 summarizes the four public management functions—and "good practice" structural design as well capacity requirements—which CAEs should be prepared to address. It also suggests indicators that evaluators can use to measure the quality of institutions governing policymaking, resource management, delivery, and intra-public sector regulation. Institutional quality—along these various dimensions—has a cumulative impact on frontline performance, including the access or reach of programs, the degree to which they are used by citizens, the extent to which they satisfy client demand, and the cost-efficiency of output production.

Getting to Good Practice: Identifying the Critical Reform Path

2.40 As described above, developments in theory and practice over the past two decades have contributed to a more comprehensive perspective on "good practice" in the design of public management systems. This notion of "good practice" should provide the basis for more positivist rather than normative or even exhortative approaches to designing and delivering assistance for institutional reform. That said, the Bank's experience with technocratic reforms of civil service and public financial management systems in developing countries suggests success requires far more than simply identifying—and filling—gaps between a country's existing institutional endowment and some "good practice." The ways in which complex interventions are sequenced and implemented have a distinct and potentially far-reaching impact on the way public management systems respond. Credible approaches to reform sequencing should therefore be based on a clear understanding of the drivers of institutional change, or the specific sequence of actions that would most likely propel public management systems toward more efficient equilibria.

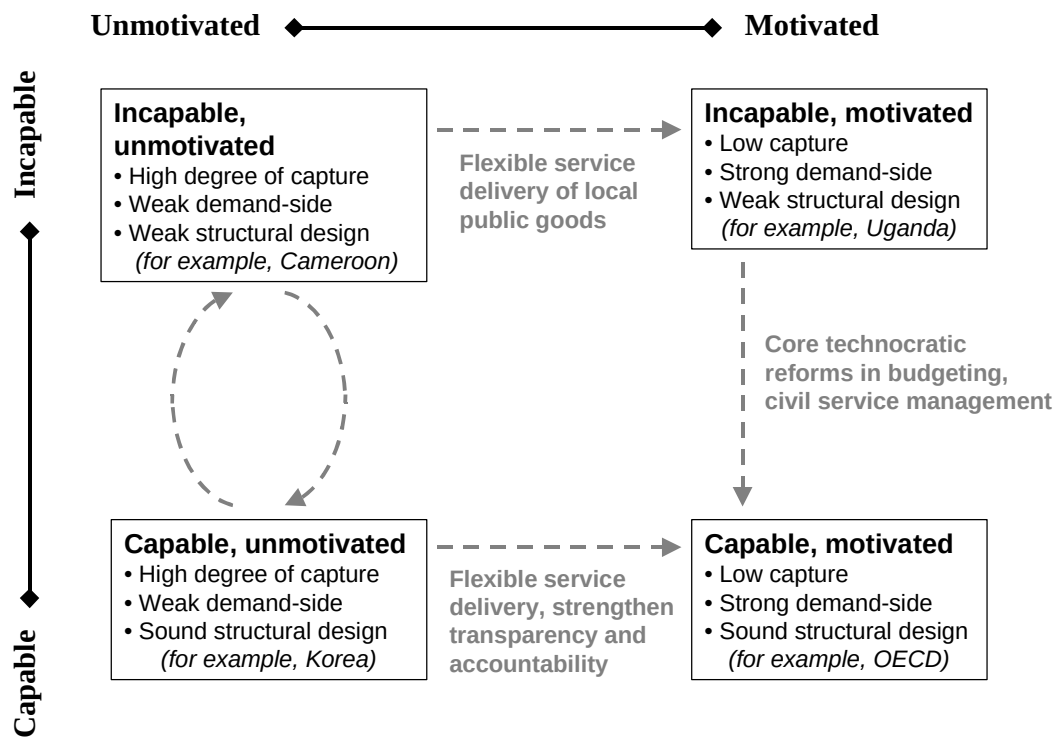
2.41 Empirical work on sequencing is rudimentary and does not offer very concrete answers for reforming governments that confront multiple structural and capacity constraints in areas such as service delivery, budgeting, personnel management, and accountability. What levers can reformers pull in order to generate virtuous cycles of long-term institutional change? Alternatively, are certain sequencing strategies (for example, de-monopolizing service delivery before strengthening core fiscal and personnel systems) more likely to precipitate vicious cycles rather than virtuous ones? How should the quality of a country's existing public management system affect its choice of sequencing strategy? While still a work in progress, the task of identifying critical reform pathways has increasingly taken center stage in the Bank's policy dialogue with a growing number of client countries. The emerging framework, described below, should help CAEs judge the specific sequencing strategies adopted by client countries over time.

2.42 The basic idea underlying this approach is that reforming governments are more likely to succeed in improving public sector *capability* (that is, the state's ability to undertake collective actions such as policymaking and budgeting, service delivery, or regulation) when they are *motivated* by demand-side pressures on the part of citizens, firms, civil society, and other stakeholders. Figure 2.7 illustrates a simple matrix based on these two stylized characteristics of public management systems, that is, motivation and capability. The matrix serves as a heuristic device not only to categorize countries, but also to think through the implications of specific

[50] Mackay, 1998.
[51] OED, 1998.

reform sequences.[52] According to the matrix, *capable and motivated* public sectors—typical of many in the OECD—are characterized by robust structures and systems for policymaking, budgeting, service delivery, and accountability. Narrow or oligarchic interests are unable to capture public institutions or public processes in part because of countervailing pressure applied by broad-based reform constituencies, organized and committed to holding government accountable. Client countries that receive Bank assistance for public sector reform are essentially seeking the critical path to this equilibrium in which indigenous demand for good governance drives a continual process of refining structure and building capacity.

Figure 2.7: Identifying the Critical Path for Public Sector Reform



Source: **Girishankar and Manning, 2000**

2.43 A growing number of states, particularly undergoing democratization and political decentralization, are increasingly facing new demand-side or motivational pressures to improve the performance of public institutions. Yet they do not possess the administrative competence to translate the preferences of their citizens into affordable policies and programs. *Motivated, yet incapable,* the public sectors in countries such as Uganda, Tanzania, and Ethiopia offer fertile soil for the implementation of technocratic reforms of planning and budgeting, as well as civil service systems. Efforts to redesign the state's administrative machinery are more likely to be scrutinized and guided by an informed public with a stake in improving government performance.

2.44 By contrast, cases of *capable, yet unmotivated* states are increasingly rare, since even ostensibly robust structures and systems have proven unstable without countervailing demand-side pressures on the executive. The process of state formation in countries such as Indonesia and

[52] Girishankar and Manning, 2000. Similar heuristic devices can be built around other stylized characteristics (for example, core public management and participatory governance) in order to highlight trade-offs that governments face when sequencing public sector reforms. Also refer to Levy, 1999.

Korea, both illustrative of the “East Asian miracle” phenomenon, resulted in equilibria characterized by administrative competence, yet captured public processes. The paroxysms of the late 1990s that followed the Asian financial crisis led to the unraveling of these arrangements and contributed to the view that capable, yet unmotivated public sectors were ultimately unstable. As predicted by the matrix in Figure 2.7, recent reform efforts in these countries have focused on transparency and accountability by maximizing opportunities for public participation in decisionmaking as well as oversight. In addition, reforms have sought to introduce flexibility-enhancing measures—primarily through decentralization and de-monopolization—in line agencies and semi-autonomous and parastatal bodies that were previously captured and vulnerable to various forms of leakage.

2.45 By and large, the Bank’s poorest client countries are burdened by public management systems, which are *incapable and unmotivated* to improve performance. Public management systems in countries such as Cameroon do not undertake collective actions to any appreciable degree on the behalf of citizens. Nor do public actors or citizens face any clear and compelling incentives to change deeply rooted patterns of rent extraction from the state. In such settings, the search for sequencing strategies that could dislodging public sectors from stable, yet dysfunctional equilibria all too often remain elusive. It is clear, however, that technocratic reforms of budgeting and civil service systems, in the absence of complementary demand-side efforts, are not sustainable, as illustrated above. Moreover, such measures may serve to reinforce dysfunctional patterns of governance by strengthening machinery without altering the underlying structure of patron-client relationships. For instance, donor-assisted efforts to downsize civil service staff within a short time horizon may actually empower the executive interested in eliminating cadres that are not obligated to them by patrimonial ties.

2.46 For countries that are lodged in the "dysfunctional equilibrium trap," the Bank’s Community-Driven Development (CDD) initiative may provide the viable entry point for a longer-term program of public sector reform.[53] Rather than start by strengthening core systems, the CDD approach would support modest resource transfers to organized communities in order to provide local public goods. Eligibility would depend on beneficiaries undertaking participatory planning, community-based oversight, and basic capacity building exercises, particularly in the rudiments of financial management and procurement. The demonstration and "learning by doing" effects of the CDD approach would be leveraged to strengthen the governance of decentralized delivery more generally, including the development of robust local government systems. The resulting pressures from communities and their representatives would then translate into greater demand for fiscal transparency and administrative competence on the part of the core public sector. The logic behind this sequencing strategy is currently the Bank-supported reform efforts in Guinea, where an adaptable program loan (APL) is being implemented to support this longer-term strategic vision for institutional change in what is a highly centralized public sector. According to one interpretation, Uganda successfully introduced systems for participatory governance during the late 1980s and early 1990s, all of which generated political demand for public sector re-orientation, including fiscal and administrative decentralization.

2.47 As noted above, this framework is based on stylized characteristics of public management systems. The same approach can be employed using other variables. The key is that CAEs provide a systematic, well-argued judgment of the sequencing strategies employed by client countries in their public sector reform efforts. Of particular importance is whether the Bank appropriately advised borrowers about the trade-offs or the risk-reward margins associated with alternative reform pathways.

[53] Girishankar and Levy, 2000.

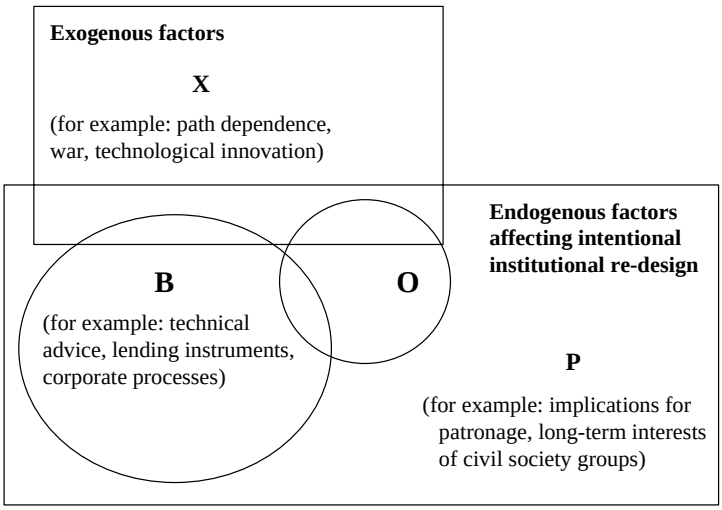
Attributing Change, Explaining Impact

2.48 Assessing Bank assistance against "good practice" approaches to structural reform, capacity building, and reform sequencing is only part of the evaluator's task. The more daunting challenge is to identify the underlying incentives and preferences that compel economic actors to influence the trajectory of institutional change and, specifically, to attribute ID impact of Bank-assisted efforts to intentionally redesign public institutions. For the purposes of CAEs, evaluators can think of ID outcomes as deriving, rather broadly, from two somewhat stylized factors—those that are *endogenous* to capacity building or structural reform efforts supported by the Bank (and other development partners) and those that are *exogenous* to the same (Figure 2.8). In attributing impact to either endogenous or exogenous factors, evaluators can draw on the rich and emerging literature on the political economy of institutional change and decisionmaking.

2.49 The Bank supports structural reform and capacity building through its dialogue with borrowers as well as its lending operations. The quality of Bank assistance (depicted as B in Figure 2.8)—including the substantive policy content of assistance, the design of operations, and the impact of Bank involvement—depends on a variety of *endogenous factors.* These include the Bank's evolving corporate strategy on public sector reform issues, its internal processes and incentives, the range of available lending and advisory instruments, and its relationships with other members of the aid community. Other actors, including other donors and borrowing governments—by virtue of their own ID initiatives (depicted as O)—significantly influence the role that the Bank plays in supporting public sector reform. These actors are also subject to various constraints imposed by their own mandates, corporate strategy and processes, etc.

2.50 Attributing the impact of Bank support requires that CAEs not only assess the substantive content of reform efforts but also analyze the political economy underpinnings (depicted as P) of ID outcomes. What did reforms imply for the interests of a particular group, specifically its ability to extract rents by capturing state institutions? Did structural reform efforts significantly alter the structure of patronage in the public sector? If so, what kinds of coalitions were required to ensure successful implementation? At what point did the benefits of public sector reform in terms of improved access to essential services or greater accountability outweigh the social costs of maintaining existing institution equilibria? Did formal political processes—including party structure, the ability of parties to facilitate bargaining between groups with divergent interests, and political influence in civil service management—shape the design or implementation of reforms? By surfacing these political economy concerns in CAEs, evaluators can not only provide richer explanations of *why* ID efforts succeed or fail, but can also identify more nuanced approaches for the Bank to facilitate public sector reforms in client countries.

Figure 2.8: Attributing Institutional Change to Endogenous and Exogenous Factors



Key: B = Interventions supported by the Bank; **O** = Interventions supported by others; **P** = Political economy factors

Source: **Author**

2.51 The trajectory of institutional change is also shaped in profound ways by *exogenous factors* unrelated to intentional efforts to reform public

institutions. Path dependencies or the cumulative effects of history have their own idiosyncratic impact on state formation. In several countries, particularly in South Asia and Africa, the colonial imprint is still evident in administrative systems, land and property rights, and state-society relations. This legacy in francophone and anglophone countries can often limit the options available to contemporary reformers of public management systems. Other exogenous factors also shape prospects for ID, including technological innovation, the development of infrastructure, economic and political geography, natural disasters, as well as war or civil strife. The key is for evaluators to distinguish, to the extent possible, between the various drivers of change and thereby develop an informed perspective on when, where, and how the Bank can make a difference in public sector reform.

2.52 This section has provided an overview of the current thinking on public institutions that matter most, their qualitative attributes, and the structural and capacity conditions needed to promote them. In addition, approaches to evaluating sequencing strategies and attributing ID impact to endogenous and exogenous variables were discussed. In the past, OED has applied some elements of this framework in evaluations of country-level ID impact. Section 3 briefly reviews a selection of past CAEs, assesses their treatment of public management issues, highlights some cross-cutting themes, and identifies lessons to guide future country-level work.

3. Public Sector Reform in Past Country Assistance Evaluations

3.01 This section assesses the treatment of public management and institutional issues in past Country Assistance Evaluations (CAEs). It seeks to identify common problems and pitfalls that evaluators face in addressing public sector reform issues, as well as key lessons that could be used to improve OED's methodology for assessing country-level ID impact.

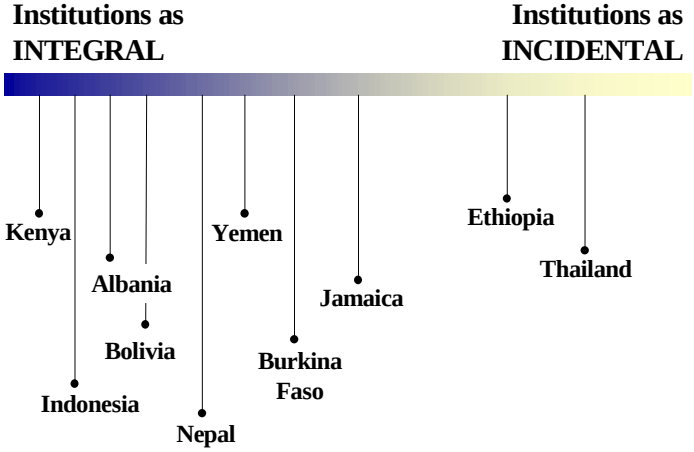
3.02 The sample for this assessment comprised 10 CAEs covering 3 AFR, 1 ECA, 2EAP, 1 SAS, 1 MNA, 2 LAC, and 1 SAS countries. The criteria for selecting these reports included geographical (or regional) distribution, experience of task managers, and level of country institutional endowment. Each report was evaluated on the basis of its coverage and diagnosis of structural and capacity constraints; the quality of its judgments regarding the relevance and efficacy of Bank assistance, and the appropriateness of its recommendations (Annex 1).

Coverage of Public Management and Institutional Issues

3.03 *Did institutions matter?* The degree to which CAEs in the sample considered institutions integral to country development prospects varied significantly. The evaluations for Albania, Bolivia, and Kenya argued convincingly that weak governance and poor capacity were among the most pressing development constraints in the near term. Evaluations of the Indonesia, Nepal, Yemen, and Burkina Faso programs also acknowledged, to varying degrees, the prominence of ID as a development concern. (Figure 3.1). These judgments were generally not supported by standard metrics for a country's overall level of institutional endowment and its implications for key indicators of growth, development, or well-being. It is worth noting that the Albania CAE concluded that development prospects were hampered by a "failed state." Such characterizations, however useful in communicating the urgency of governance problems, did not constitute a sufficiently rigorous benchmark against which to assess governance improvements in the future.

3.04 The Ethiopia and Thailand CAEs were exceptions. In both cases, the structure of governance institutions was considered incidental to country development prospects. To the extent that it addressed these issues, the Thailand CAE viewed ID impact as primarily contingent on the Bank identifying well-intentioned bureaucrats to serve as interlocutors and implementing agents. At any rate, neither report adequately captured the increasingly sophisticated dialogue that the Bank has managed to develop with both its Ethiopian and Thai clients in recent years. For instance, in 1998, the Government of Ethiopia (GOE) drafted a Capacity Building Strategy and Programme Framework for mitigating capacity constraints in federal government, the regional governments, and the private sector, all of which hinder the implementation of key policies and programs designed to reduce poverty. On the request of the GOE, the Bank has been engaged in a vigorous dialogue—including the development of an ambitious analytical agenda—

Figure 3.1: How Did CAEs View Institutions?



on the structural underpinnings of effective public sector capacity building. For their part, the Thai authorities have developed a comprehensive public sector reform covering issues of public financial management and administrative reform with the support of a Bank-financed programmatic SAL.[57] Neither initiative was addressed in these CAEs.

3.05 *Which institutions were critical?* A central challenge for CAEs—and evaluation in general—is to move from largely axiomatic claims about the importance of institutions to more systematic analyses of institutional quality and its implications for various aspects of wellbeing such as access of the poor to essential services, their consumption patterns, and degree of empowerment. Using the framework described in section 2, this review assessed how effectively CAEs diagnosed the quality of public management systems—specifically, structural and capacity constraints on policymaking, resource management, service delivery, and regulation.

3.06 As Table 3.1 illustrates, CAEs diagnosed public management systems in most sample countries as being burdened by insulated sectoral policymaking processes (Kenya, Indonesia, Jamaica, Nepal) and centralized, monopolistic delivery systems (Albania, Burkina Faso, Yemen). Both these structural constraints undermined the allocative and operational efficiency of public expenditures. Personnel management problems—specifically, poorly paid yet overstaffed civil service cadres—were endemic. In addition, even when countries made planning processes more inclusive and delivery arrangements more competitive, they experienced various degrees of budgetary uncertainty in part because of problems of political interference and corruption (Bolivia, Kenya, and Indonesia and, to a lesser extent, Burkina, Nepal, Yemen). Absent credible and enforce regulation, arbitrary actions by public actors went unchecked (Nepal).

Table 3.1: Structural and Capacity Constraints on Transformation Diagnosed in CAEs

CAE	POLICY		RESOURCES		DELIVERY		REGULATION		
	MACRO	SECTORAL	BUDGET	PERSONNEL	DECENT.	DEMONOP.	INPUT CTRLS.	BENCHMARKS	M&E
Albania	**C**	**C**		**S**	**SC**	**S**			
Bolivia	**C**	**SC**	**C**				**SC**		
Burkina Faso		**SC**	**SC**	**S**	**SC**	**SC**	**SC**		**SC**
Ethiopia						**S**			
Indonesia		**S**	**C**	**SC**	**SC**	**SC**	**SC**	**S**	
Jamaica	**C**	**S**		**SC**		**C**			
Kenya		**S**	**S**	**SC**			**SC**		
Nepal		**S**		**S**		**SC**	**SC**		
Thailand						**S**			
Yemen				**SC**	**C**	**SC**	**SC**		

where S=structural constraints and C=capacity constraints

3.07 On the capacity side of institutional performance, skills gaps were manifest in a variety of areas—from economic management in Albania to basic literacy among newly elected local officials in Burkina. Some evaluations found that poor capacity in accounting, bookkeeping, and other aspects of inputs management not only weakened budget execution, but program implementation in general (Kenya, Bolivia). Overall, CAEs could not provide a clear perspective on systemic change in part because their coverage and analysis of critical issues such as budgeting, line operations, and intra-public sector regulation were wanting (Figure 3.1).

[57] Africa Capacity Building Group, 1999; Thailand Public Sector Reform, Presentation to PREM, 1999.

3.08 *Weak attributions, limited use of data.* Most CAEs attributed sectoral outcomes (such as literacy or health status) to changes in the access and, to a lesser extent, the utilization of basic services. Accordingly, the reports used data on access (for example, percentage of population with access to safe drinking water) and usage (enrollment rates in primary education) to illustrate trends in the *reach* of basic services. Other dimensions of performance such as client satisfaction, efficiency, or leakage and their impact on sectoral outcomes were not adequately explored. The only exception was the Kenya report, which noted that cost inefficiencies and leakage in facilities had adversely affected health outcomes.[58]

3.09 The core concern of ID impact evaluation—attributing frontline performance to institutional design—was left largely untouched in CAEs. In fact, much of the quantitative or perception data on institutional quality (for example, budget variance, skills utilization, rule of law, and incidence of corruption) required to make such attributions was absent in these analyses. Even measures of real wages, decompression ratios, and private-public sector salary differentials—long recognized as key indicators of state capability—were lacking. Only in the Yemen CAE were estimates of wage differentials between Project Implementation Unit staff and other civil servants provided. It was more common for the evaluations (Burkina, Kenya, Yemen) to report the fiscal impact of various approaches to civil service pay and employment (for example, the wage bill as a percentage of recurrent expenditures); the Albania CAE also analyzed civil service staffing trends in arguing for tighter establishment control.

Relevance of Bank Strategy

3.10 Most of the CAEs in the sample concluded that the Bank lacked a coherent strategy on public sector reform or country-level ID, even when sectoral ID objectives were deemed relevant. In evaluating relevance, the reports did not adequately differentiate between structural and capacity constraints. Most CAEs, however, did go beyond the stated objectives of projects and considered whether the relevance of country programs was preserved in the details of operational design, and their "goodness of fit" to institutional setting. Some evaluations in turn attributed flaws in both strategy and design to internal disincentives within the Bank itself.

3.11 *Disconnect between sectoral and country-level ID strategies*. Nearly half the CAEs in the sample found that the Bank failed to articulate a coherent public sector reform or ID strategy in its country dialogue (Albania, Burkina, Kenya, Indonesia, Yemen). At the same time, the reports judged ID objectives within specific sectors or specific operations to be relevant. For example, a Bank-financed social fund that supported bottom approaches to micro-lending in Albania was deemed highly relevant, given the urgent need to decentralize and de-monopolize rural credit institutions. Similarly, Bank assistance for decentralized management of irrigation systems was also viewed favorably.

3.12 Some CAEs attempted an explanation of this apparent *disconnect* between relevant sectoral ID strategies and weak country-level ones. First, eliciting commitment for comprehensive reform efforts was considered more difficult than gaining support for less ambitious, sector-specific interventions. Second, in countries with weak public management

[58] The frontline performance data required to make these attributions are not always reliable. To the extent possible, evaluators should rely on household survey data undertaken in the context of Poverty Assessments to collect access, utilization, and client satisfaction data for essential services such as education, health, water, and roads. In addition, they can draw on Public Expenditure Reviews (PERs) that calculate cost-efficiency data using aggregate expenditure and output data. Finally, evaluators may also from data generated by expenditure tracking as well as facility-level, cost-efficiency surveys—both of which have gained currency as instruments for measuring actual levels of leakage as well as output performance.

systems yet rapidly expanding portfolios, the Bank faced an incentive to continue with limited capacity building efforts—channeled through Project Implementation Units (PIUs)—to facilitate the implementation of new projects and, therefore, make immediate, demonstrable improvements in wellbeing. For their part, the Bank's interlocutors—namely, senior civil servants and politicians—generally have an interest in maintaining office and are therefore more likely to favor opportunistic ID interventions with immediate pay-offs rather than systemic reforms with long-term pay-offs. The projectized approach to ID taken by the Bank in its Albania and Yemen programs was indicative of this incrementalist bias. Notably, the Yemen CAE, and to a lesser extent, the Albania CAE, critiqued the Bank's strategic choice, arguing that sectoral interventions using enclaves (PIUs or social funds) promoted "islands" of capability even as larger systems of public management deteriorated. Although insufficiently highlighted in CAEs, this critique of the projectized, incrementalist approach to ID implies that Bank support for more comprehensive reform would require greater reliance on programmatic lending.

3.13 *No differentiation between structural and capacity concerns.* For the most part, evaluations of relevance tended to conflate structural and capacity aspects of institutional dysfunction. The reports typically evaluated whether the Bank supported efforts to remedy "weak public administration," "poor institutional capacity," and "weak governance." Absent a framework for distinguishing between structural and capacity constraints, OED was unable to incorporate positivist analyses of ID strategy and sequencing in evaluations of relevance.

3.14 *Design and degree of fit considered significant.* In evaluating relevance, CAEs typically examined whether operations—once designed—were appropriately aligned with Bank strategy and also fitted to institutional context. CAEs across countries concluded that the design of Bank operations was often too complex and ambitious for the prevailing institutional environment. In addition, the Bank tended to overestimate the capacity of counterparts (Jamaica, Nepal, Burkina, Albania, Bolivia) and the quality of government control systems to ensure compliance with procurement and financial management standards (Kenya). Similar concerns for weak regulatory and control systems were not evident in CAEs' analyses of Bank-financed social funds, AGETIPs, road funds, and other semi-autonomous institutions (Albania, Burkina, Bolivia, Ethiopia, Yemen). This lacuna is worth noting since Bank support for institutions with arms-length relationships to the core and, at times, off-budget financing—absent complementary efforts to strengthen intra-public sector regulation—illustrates how the incrementalist approach, discussed in para 3.12, can inadvertently weak the public sector fiduciary framework.

3.15 According to most CAEs in the sample, poor operational design often resulted from overly technocratic approaches to Economic and Sector Work (ESW) as well as project identification and preparation. For example, the failure to involve government more directly in ESW led to poor ownership of analyses and recommendations (Kenya, Indonesia, Nepal, Yemen). Lack of participation during project design also obscured ground realities, local capacities, and local preferences (Nepal). These evaluations in turn recommended greater participation earlier in the project cycle, as well as greater linkages between ESW and project design in order to improve the degree of fit to the local institutional context (Albania, Kenya, Nepal).

3.16 Relevance was also undermined when the Bank failed to respond swiftly to emerging crises either by developing new operations or restructuring existing ones (Albania, Indonesia, Thailand). Some CAEs attributed poor responsiveness to the organizational cultural within country teams. For instance, in Albania and Indonesia, evaluations argued that a "halo effect" within the Bank reinforced over-optimistic ratings of portfolio performance and projections of development prospects. These views persisted in spite of documented warnings about increasing corruption, weakening rule of law, and unraveling political equilibria.

Efficacy of Bank Support

3.17 *Problems in "scaling up" ID impact from the project to country level.* Most CAEs reported on the impact of ID efforts at the sectoral and project levels. Nearly half found that operations, particularly in service delivery sectors such as infrastructure, social sectors, and agriculture, resulted in improved performance (Albania, Bolivia, Burkina Faso, Ethiopia, Indonesia). In the remaining reports, the ID impact was hampered by weak capacity, poor accountability, and lack of sustainability (Jamaica, Nepal, Yemen). These findings were typically substantiated using ID impact ratings from OED project audits rather than actual frontline data (with modest exceptions in the case of the Kenya and Nepal CAEs).

3.18 Despite some of the sectoral or project-level gains noted above, nearly half the CAEs found the Bank's approach to be ineffective *at the country level* (Albania, Bolivia, Indonesia, Kenya, Yemen). None of these CAEs, however, offered a systematic methodology for "scaling up" OED's project-level ID impact ratings to the country level (Bolivia, Burkina Faso, Kenya, Indonesia, Nepal). Simple non-weighted averages of project-level ratings from each portfolio were used as indicators or country-level ID impact. This approach was unable *inter alia* to account for the differential impact on institutional endowment of distinct reforms (for example, the impact of improving civil service pay on public sector performance as compared to that of de-monopolizing urban infrastructure delivery). It is worth noting that the Kenya CAE did use the Transparency International Corruption Index to illustrate the decidedly modest contribution of Bank assistance in forestalling further decline in public sector institutional quality. While they are helpful in gauging the *credibility* of state institutions, such indices are based largely on epiphenomenal survey data, which are subject to lags as well as moral hazard. These by themselves are not adequate proxies for measuring the impact of various Bank-supported interventions on the public sector transformation process as a whole (Annex 2).

3.19 As a rule, evaluations of efficacy require an appropriate counterfactual—one that helps determine the added value of Bank support to improvements in country institutional endowment. Most CAEs did not delineate and test counterfactuals in terms of institutional performance benchmarks, either for a given country over time or against country comparators. One exception was the Yemen CAE, which defined a counterfactual to evaluate—in qualitative terms—the widespread use of PIUs in response to poor capacity on the ground. The report concluded that Bank support for core public sector reform—rather than creation of new enclaves—would have resulted in greater country-level ID impact.

3.20 *Endogenous factors.* Most CAEs attributed poor implementation and efficacy of ID to three factors that were endogenous to Bank-supported interventions. These were the quality of corporate processes (within the Bank), political commitment to reform, and the unintended consequences of enclavist approaches to ID. As far as *corporate processes* were concerned, CAEs noted that high rates of staff turnover slowed implementation of ongoing ID efforts. New staff typically lacked sufficient knowledge of the country's institutional context and of the key players on the ground (Albania, Burkina Faso, Kenya). Second, poor country knowledge was in part the result of under-investing in ESW on governance (Ethiopia, Albania). Even when ESW was undertaken, it was not sufficiently rigorous, often sequenced after operations became effective (Albania), or simply not incorporated in project preparation. In some cases, CAEs found that staff were not even aware that ESW had been conducted (Burkina).

3.21 A third set of corporate processes affecting the efficacy of ID was supervision and monitoring. The Indonesia and Jamaica CAEs found the intensity of supervision and M&E to be insufficient given the need for close monitoring of institutional change dynamics on the ground.

According to the Indonesia report, the quality of supervision of ID efforts suffered because of the "halo effect" mentioned above. The remaining evaluations did not adequately address M&E concerns. Fourth and finally, CAEs found that weak coordination between donors led to duplication and improper sequencing of reforms such as downsizing and functional analysis of core ministries (Albania, Nepal).

3.22 CAEs found that sound technocratic knowledge of institutional constraints and well-designed ID interventions were not sufficient in promoting public sector performance (Albania, Jamaica, Indonesia). In other words, politics mattered. Aside from noting its importance, CAEs did not offer a more sophisticated approach to the policy economy of institutional reform. The reports overlooked *inter alia* the role of interest groups and formal processes that constrain potential "champions" of reform. Nor did they explicitly analyze how the incentives that senior decisionmakers face may contribute to the under-performance of more ambitious ID interventions. Sequencing reforms of different public management functions (for example, core budgeting and administrative reform vis-à-vis decentralized delivery) was also not addressed.

3.23 Three evaluations in the sample did highlight the *unintended consequences* of efforts to improve performance by circumventing established—albeit weak—public sector institutions (Albania, Bolivia, Yemen). These efforts typically involved the creation of financially endowed PIUs or topping up of civil service salaries to facilitate project implementation. The Bolivia CAE argued that attempts to use salary top-ups ran counter to ongoing civil service reform efforts to improve the overall incentive environment. In Albania and Yemen, CAEs concluded that the Bank repeatedly underestimated the opportunity costs (of creating semi-autonomous units to speed up project implementation) in terms of forgone development of the core public sector. It is also worth noting that the CAEs could not verify that these short-term efforts in fact reduced capacity and skills gaps. To the contrary, fast growing portfolios in Albania and Bolivia likely imposed even greater capacity requirements on already weak state institutions.

3.24 *Exogenous factors affect counterfactuals.* Exogenous factors were typically given less weight in explaining outcomes than endogenous ones. Nevertheless, two such exogenous factors were salient in borrower efforts to implement public sector reform and ID initiatives. First, existing weaknesses in the capacity and structure of state institutions slowed (and even stalled) the successful implementation of ID efforts. These typically involved high turnover of staff in implementing agencies (Nepal) as well as political interference during implementation (Kenya).

3.25 A second, broader set of constraints comprised path dependencies that shaped the nature of state-society relations. Whether difficult political transitions (for example, from an autocratic to democratic regime in Albania) or efforts to extend the reach of legitimate authority (for example, to include semi-nomadic tribes in Yemen), borrowers inherited significant constraints to formal rule-based governance (Albania, Indonesia, Kenya, Yemen). At times, the paroxysms associated with such historical processes such as Albania's 1997 civil conflict severely undermined any attempt to improve public management systems or undertake even incremental ID initiatives.

Quality of Recommendations

3.26 CAEs recommendations were generally consistent with evaluation findings, stressing the need for a more systemic focus in addition to more idiosyncratic interventions that could deliver tangible improvements in frontline performance. This implies the need for a more convincing, systematic approach—which the CAEs did not provide—to sequencing short-term and longer-term reform initiatives in ways that are political feasible. The evaluations did, however, stress the need for building actionable, timely knowledge on institutions.

3.27 *Recommendations consistent with findings.* CAE recommendations on public sector, institutional issues did focus on the deficiencies highlighted in evaluation findings (Albania, Bolivia, Burkina Faso, Indonesia, Kenya, Yemen). These reports recommended that the Bank focus future ID efforts in priority areas that directly affected frontline delivery performance. These included areas such as decentralized service delivery (Bolivia, Kenya, Nepal), capacity building (Bolivia, Burkina Faso), and accountability (Kenya). Notably, in four countries, evaluations drew a distinction between structural reform priorities (for example, improving public sector oversight mechanisms in Indonesia, Kenya) and capacity building ones (that is, increasing resource allocation in Burkina Faso or capacity in Bolivia). In Yemen and Albania, where the Bank's ID strategy was largely enclavist, CAEs recommended a *systems* approach to reinforce rather than circumvent the core public sector. That said, recommendations in the Ethiopia, Jamaica, and Thailand CAEs were too generic and superficial to be of operational value. For instance, the Ethiopia report recommended that the Bank "promote ID at all levels." In Jamaica, where the Bank had supported serial civil service reform operations, the CAE only argued for more vigorous support for "institutional strengthening."

3.28 *Weak on entry points and sequencing strategies.* Nearly half the CAEs cited borrowers' political commitment—rather than inadequate knowledge of institutional context—as the key obstacle to implementation of recommended institutional reforms (Bolivia, Kenya, Nepal, Yemen). Yet none of the evaluations identified politically feasible "entry points" for reform. Nor did they elaborate on sequencing strategies that build on demonstrable results in the short run, while ensuring that the trajectory of ID follows a critical path to long-term reform objectives.

3.29 *The need for actionable knowledge.* Knowledge of institutions in and of themselves was not sufficient to carry out institutional reforms (Bolivia, Kenya, Yemen). To be actionable, knowledge of institutions should be produced strategically and on a timely basis. In other words, analytical work on institutions should be more tightly incorporated in the work and decisionmaking processes within operational units. Partly in recognition of this need, the Albania and Indonesia CAEs argued for country-monitoring systems to alert Bank senior management about rapidly changing institutional and political dynamics. Others also stressed the need for more intensive monitoring and evaluation of Bank-supported ID efforts (Burkina Faso, Yemen).

Findings Summarized, Lessons Learned

3.30 In addition to providing a benchmark for future CAEs, the following findings and lessons should help developed a more standardized approach to country-level evaluations of Bank-assisted public sector reform and ID (see Section 4):

3.31 *Mapping constraints helps gauge responsiveness:* The quality of CAEs was weakened by lack of systematic "mapping" of structural and capacity constraints on key public management functions such as policymaking, resource management, delivery, and regulation. In addition to helping pinpoint the location and severity of ID needs, "constraints maps" can allow evaluators to judge the responsiveness and relevance of Bank-supported interventions (Annex 2).

3.32 *Impact assessments require data on institutional quality and frontline performance:* CAEs did not use data effectively in assessing the impact of Bank support on either the quality of key public institutions (such as those governing budgeting or service delivery) or frontline performance. This in turn hampered the ability of evaluators to effectively delineate counterfactuals for Bank support of ID. To improve the use of data in the short run, CAEs should rely on available data on fiscal transparency, budget variance, and other observable aspects of institutional quality. In addition, household survey data developed in the context of Poverty

Assessments should be mined more thoroughly for frontline performance indicators on access, coverage, and satisfaction. Over time, CAEs should systematically rate the institutional quality of public management systems (using a questionnaire and score card such as those one provided in Annex 2) in order to set standardized benchmarks, which can then be used in future CAEs.

3.33 *Counterfactuals are more robust when exogenous factors are considered:* CAEs generally analyzed how internal processes such as staff turnover, supervision intensity, and even organizational culture affected the Bank's ability to effectively support ID. They were less rigorous in analyses of exogenous factors such as state-society relations. Recent advancements in constructing indicators on governance quality should provide an empirical base for incorporating exogenous factors when defining counterfactuals.

3.34 *Political economy analyses are central to sequencing:* Especially worthy of note is the importance of mainstreaming the political economy of institutional change as an evaluative concern in CAEs. As discussed in Section 2, credible approaches to sequencing require equally sophisticated explanations of the political economy underpinnings of institutional equilibria and the drivers of change in the public sector.

4. Evaluating Bank Support for Public Sector Reform

4.01 Based on the current thinking on public management, as well as lessons from the assessment in Section 3, this section offers a step-by-step approach for evaluating the country-level impact of Bank-supported public sector reform and ID efforts. The approach proposed here conforms to OED's established evaluative concerns, namely, the relevance and efficacy of Bank support. It also proposes that CAEs evaluate Bank "responsiveness" to institutional constraints. Each step is explained below with illustrative examples. In addition, Annex 2 presents a sample questionnaire and score card for evaluators currently undertaking CAEs.

Responsiveness and Relevance

4.02 Evaluators should determine whether the Bank articulated a discernible country-level ID strategy. In addition, they should judge the *responsiveness* of specific Bank-supported interventions to capacity and structural constraints on public management. Measuring "responsiveness" allows evaluators to systematically incorporate issues of design into their overall judgment of relevance.[59] After mapping the location and severity of key public management constraints, evaluators should be able to judge whether the Bank supported ID along critical and *relevant* reform pathways. The impact of upstream factors on responsiveness and relevance should be examined; these include internal incentives facing Bank task team leaders and managers, choice of lending instruments, commitment of the borrower, and involvement of other donors in public sector reform and institutional issues.

4.03 Step 1: Score and map the severity of public management constraints. A first step for country-level evaluations of ID impact should be a systematic diagnosis of initial institutional constraints on public management systems. Structural and capacity constraints prevalent in specific sector or tiers of government can be rated and aggregated along the four critical axes of transformation—policymaking, resource management, delivery, and regulatory enforcement. Evaluators should also gauge how structural and capacity constraints (for example, insular sectoral planning processes) affected relevant aspects of *institutional quality* (for example, policy volatility and incoherence). Annex 2 provides a detailed questionnaire for mapping and scoring structural and capacity constraints on a six-point scale (0-5) and for benchmarking their impact on institutional quality.

4.04 For sake of illustration, Figure 4.1 reproduces the "constraints map" for a public management system that is *structurally constrained* in resource management and regulation, and *capacity-constrained* in policymaking and delivery (Annex 2). The "radar chart" shows that structural reforms would be required for the "inputs regime" (or the link between policymaking and resource allocation functions) as well as "output production" (or the translation of inputs into outputs). Similarly, capacity building needs are apparent not only for organizations involved in output production but also those entrusted with monitoring the impact of service delivery on development outcomes.

[59] Integrating design in judgments about relevance enables OED to determine whether Bank commitments in fact went beyond broad statements of intent.

Figure 4.1: Scoring and Mapping the Severity of Initial Structural and Capacity Constraints

	Policy	Resource	Delivery	Regulation
Structural constraints	9.00	4.00	6.00	2.00
Capacity constraints	5.00	6.00	7.00	3.00

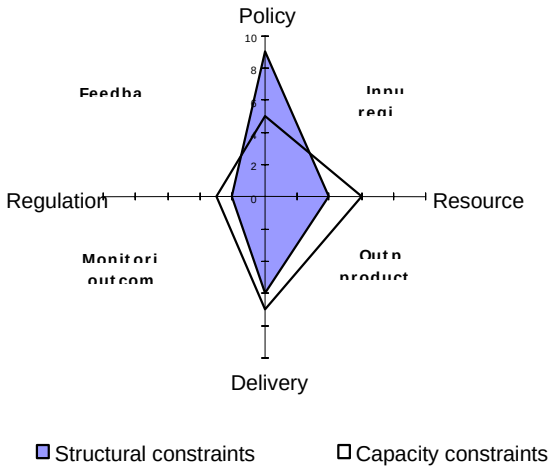


Figure 4.2: Scoring and Mapping the Intensity of Structural Reform and Capacity Building Interventions

	Policy	Resource	Delivery	Regulation
Structural reform	10.00	0.00	0.00	0.00
Capacity building	0.00	10.00	10.00	10.00

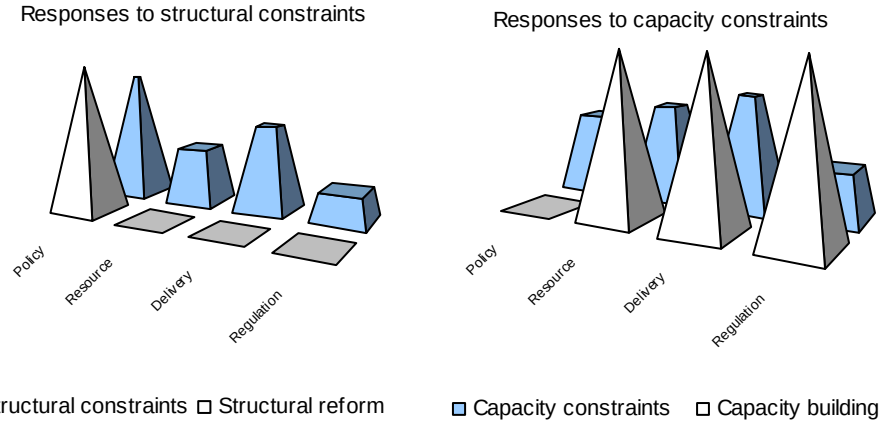
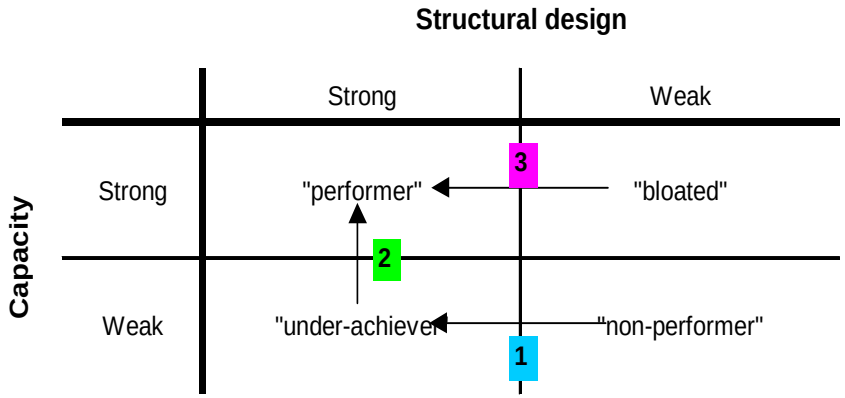


Figure 4.3: Sequencing Interventions Based on a Typology of Organizations



Path	Reform measures
1	***"Light" structural reform:*** Strengthening checks and balances including competition and voice; and disseminating basic information to elicit demand-side response from clients
2	***Capacity building:*** Investing in plant and equipment; supporting skills development at the individual and institutional-levels; more competitive mechanisms for delivering training
3	***"Heavy" structural reform:*** Enforcing hard budget constraint; rationalizing fiscal, pay, and employment systems; developing more knowledge-intensive information systems; strengthening oversight mechanisms based on public reporting of financial and operational audit information, expanding options for competition and voice.

4.05 Step 2: Categorize and score the intensity of Bank support. Once initial constraints have been mapped, evaluators should *categorize* Bank-supported interventions in terms of policymaking, resource management, delivery, or regulation. Just as they diagnosed constraints, evaluators should rate and score *intensity* of interventions in supporting structural reform and capacity building on a six-point (0-5) scale. Subsequently, maps of constraints and interventions should be superimposed so evaluators can assess the *responsiveness* of Bank support. Figure 4.2 illustrates the case in which the Bank—confronted with the "constraints map" noted in Step 1—responded with a set of interventions that over-emphasized capacity building, while ignoring structural reform of budgetary, delivery, and regulatory institutions. In this case, evaluators could reasonably judge Bank efforts to be only marginally responsive. To substantiate this judgment further, "response rates"—the ratio of intervention intensity to constraint severity—can be calculated either for specific functions such as service delivery or for public management systems as a whole (para. 11, Annex 2).

4.06 Step 3. Chart critical pathways, judge relevance. To judge relevance credibly, evaluators require a clear understanding of the critical path to a modern system of public management, as well as an appropriate strategy for sequencing efforts to improve distinct elements of public sector transformation (for example, whether core budgeting reforms should be undertaken before decentralized service delivery). Section 2 provide a framework for thinking through strategic sequencing for comprehensive public sector reform efforts. By locating public sectors within the *capability motivation* typology presented in Figure 2.8, evaluators should be able to identify more desirable pathways for long-term systemic change (for example, a focus on decentralized delivery of local public goods rather than core technocratic reforms for incapable and unmotivated states). Various sources of data including governance indicators being prepared by the Poverty Reduction and Economic Management (PREM) Network and the World Bank Institute (WBI) can help gauge a public sector's level of motivation and capability.

4.07 In addition to identifying the critical path for macro-level, systemic change of the public sector, evaluators should be prepared to evaluate Bank assistance to organizations involved in carrying out specific functions (such as the Ministry of Finance or a social fund involved in service delivery). Using data generated by the questionnaire presented in Annex 2, evaluators can categorize specific agencies or institutions in terms of the quality of their structural design and capacity. The structural design-capacity matrix in Figure 4.3 enables evaluators to categorize specific organizations that received Bank support as non-performers, bloated organizations, underachievers, and performers.[60] For each category of organization, evaluators can use the matrix in Figure 4.3 to identify more desirable pathways for sequencing structural reforms and capacity building activities. The three pathways highlighted in Figure 4.3 draw on a key principle: ID efforts are most effective at the organizational level when reforms of institutional structure precede infusions of capacity inputs such as skills and equipment.

4.08 Accordingly, supplying skills and equipment to *non-performers* is not likely to result in sustainable improvements in the organization's capacity or structural design. Rather, organizations that are structurally weak and capacity-constrained would benefit from "light structural reform" efforts designed to generate external pressures on employees to perform. These include participatory feedback from clients, options for them to opt out and receive services from other institutions, and establishment of checks and balances such as regular monitoring. Since these types of organizations have very little capacity, they probably will not be able to implement more substantial technocratic reforms of pay and internal incentive systems

[60] Placing a set of institutions along this performance spectrum requires some arbitrary judgment about the range of severity scores that corresponds to each category (for example, institutions in the bottom quartile, with scores of 0-1.25, could be considered non-performers.

immediately. Over time, external pressure to perform would create an environment in which employees would develop skills and efficiently using existing assets. More ambitious structural reforms would be appropriate. These in turn would demand more of employees in terms of cognitive skill and complementary inputs, at which point further capacity building would generate high returns. Many service delivery agencies, particularly in the social sectors, would respond favorably to the sequence of interventions described above. This is in part because clients would be willing to apply demand-side pressures on non-performers.

4.09 *Underachievers* are organizations that satisfy the basic structural conditions for performance. Their binding constraints are know-how, the capacity of technology, lack of equipment, etc. These institutions would likely make demonstrable gains in performance merely by traveling along path 2, which constitutes, for all intents and purposes, traditional capacity building. There is a third group of problem institutions—what we may call *bloated organizations.* Typically, they are overstaffed, usually in arrears, and yet they continue to under-utilize the inputs and know-how are at their disposal. Path 3 is appropriate for such organizations. It involves the imposition of a hard budget constraint as well as "heavier structural reforms" such as traditional pay and employment restructuring, developing more knowledge-intensive information systems, and strengthening the more sophisticated oversight and monitoring mechanisms. Oftentimes, core economic agencies such as Ministries of Finance, central banks, and cabinet offices are good candidates for path 3 interventions.

4.10 Typologies such as those described above tend to conflate complex institutional dynamics. Nevertheless, these heuristic devices allow evaluators to chart critical institutional change pathways—and appropriate sequencing strategies—both at the systemic and the organizational or operational levels. As mentioned earlier, past OED evaluations found that the Bank historically supported capacity building efforts for non-performers even before they undertook structural reforms. In addition, Bank-assisted civil service reforms across regions imposed hard budget constraints on personnel expenditures in core ministries without simultaneously strengthening monitoring and enforcement of establishment levels.[61] Using the approach described above, evaluators should be able to more rigorously investigate and assess these types of sequencing problems, which typically undermine the relevance of Bank assistance.

4.11 Step 4: Identify upstream factors shaping responsiveness and relevance. In explaining empirical findings on relevance, evaluators should consider upstream variables that are both endogenous and exogenous to ID efforts. As noted in Sections 2 and 3, factors influencing responsiveness and relevance include the Bank's own corporate processes, internal incentives facing Bank task team leaders and their managers, political incentives facing senior government officials, the role of other donors, and exogenous factors such as state-society relations.

4.12 *Upstream corporate processes,* such as the quality of ESW or other forms of institutional analysis, can significantly influence the Bank's decision to raise public sector reform issues in the dialogue with its clients. This type of analytical work also influences the design, and therefore the responsiveness, of specific operations. Equally important are the quality of Bank staff leading the public sector reform dialogue, their turnover rates, and incentives within the country teams to address public sector reform and institutional issues.[62] The choice of lending instruments also shapes responsiveness and relevance. For instance, quick-disbursing adjustment lending—while effective in supporting "stroke of the pen" reforms—has proven less effective in providing

[61] OED, 1999a.

[62] Planning and Budgeting Department (PBD) data on staffweek commitments during identification, pre-appraisal, and appraisal are generally not disaggregated by ID components in specific operations. However, such data can be collected for public sector reform operations and compared on a cross-country basis.

sufficient flexibility for far-reaching institutional reform efforts.[63] In recent years, the Bank has begun to use a wider range of lending instruments to support institutional reform such as long-range Adaptable Program Loans (APLs) in Tanzania, Guinea, and Ghana; multi-sectoral budget support operations called Poverty Reduction Support Credits (PRSCs) in Uganda; and programmatic Structural Adjustment Loans (P-SALs) in Thailand. Greater choice of instruments implies that operational staff, and therefore evaluators, will need to pay closer attention to issues of design such as the degree of flexibility in managing project funds, the importance of performance benchmarking, and the nature of conditionality (for example, inputs-based versus outputs-based).

4.13 *Political economy* factors mentioned in Section 2 can be used to unbundle "borrower ownership." CAEs should provide a much richer political economy perspective on whether the design of Bank assistance was supported by conditions favorable to public sector reform. For instance, were proposed reforms supported by a stable coalition of actors capable of resisting pressures for policy reversal? Were reform proposals discussed and debated through the formal processes of parliamentary politics? Was there a critical mass of political actors from different parties or factions prepared to agree to reductions in the aggregate level of patronage in civil service management? If so, how could their involvement in the design of reforms be further consolidated? If not, were civil society or other groups with "lower discount rates" systematically involved at the design, implementation, and monitoring of reforms in order to temper the short-term interests of senior civil servants and politicians.[64] Bringing these and other political economy to the surface is critical to understanding *why* certain approaches to external assistance tend to work while others do not.

4.14 A third consideration is *coordination among donors.* When other donors have had a comparative advantage in supporting public sector reform or ID initiatives, it would be more effective for the Bank to play a supporting role. In such cases, the Bank's country-level ID strategy may legitimately focus on leveraging analytical resources through strategic partnerships with lead donors. To make such judgments, evaluators should be cognizant of the comparative advantage and core competencies of various donors, which may vary by region and country.

4.15 Finally, Bank support for ID is also conditioned by the bargaining position of clients themselves. It is generally more difficult to address critical public sector reform issues—and especially those related to corruption and patronage—in the dialogue with large borrowers. At the same time, the Bank is better positioned to address such issues in smaller, more aid-dependent countries. The proclivity of senior decisionmakers in client countries to support far-reaching reforms of the state is also shaped by historical factors that are wholly *exogenous* to donor assistance (for example, the state's relationship with traditional leaders and institutions, ethnic conflicts with extra-territorial implications, geo-political dynamics). These provide a backdrop to CAE assessments of the relevance of Bank support.

Efficacy and Impact

4.16 Credible assessments of efficacy and impact require that evaluators have a systematic approach to addressing problems of aggregation and attribution. OED's standard methodology asks evaluators to judge interventions as efficacious if they met their stated or intended objectives. Since CAEs purport to "move to a higher plane" to assess the *country-level* impact effect of myriad ID efforts, it is not sufficient to look only at stated objectives of interventions for two reasons. First, stated objectives may themselves run counter to improving public

[63] OED, 1999a.
[64] Sheplse, 1998; Haggard and McCubbins, 1997; Johnson and Libecap, 1994.

management systems as a whole. For instance, in some countries, facilitation of PIUs and other enclaves has weakened morale of civil service cadres as well as the credibility of personnel management rules that should govern the civil service. ID interventions may also have numerous *unintended* effects on the public sector transformation process. For this reason, evaluations of impact need to capture the cumulative impact of structural reform and capacity building interventions on a country's public management system.

4.17 The second issue concerns attributing changes in institutional quality and frontline performance to Bank-supported ID interventions. Such changes can result from any number of factors, including the design of Bank operations, the degree of information asymmetry (and therefore moral hazard or adverse selection) during implementation, supervision and monitoring, or simply the "accidents of history." In explaining impact, evaluators should be prepared to make such attributions systematically and, if possible, empirically. The suggested Steps 5-7 seek to limit problems of aggregation and attribution and help evaluators assess the efficacy of Bank support and identify the factors shaping impact on the ground.

4.18 Step 5: Collect data on changes in the transformation process. ID efforts affect frontline performance by first alleviating *structural and capacity constraints* and, in the process, improve *institutional quality*. This in turn provides conditions favorable to improved output performance (table 2.1). The questionnaire and scorecard in Annex 2 offer one approach to benchmarking structural, capacity, and institutional quality variables over time or against comparators.

4.19 Until OED has developed a robust database with these institutional and public management indicators, evaluators can consult with PREM's cross-country database on governance-related indicators. In addition to supporting the implementation of various survey-based institutional analyses across client countries, PREM is developing a scorecard that draws data (some of which include structural and capacity constraints) from multiple sources to track changes in institutional quality.[65] WBI has also begun implementing triangulated surveys of public officials, private firms, and beneficiaries in several countries. These "governance" surveys seek to develop benchmarks for various capacity and structural constraints (for example, the level of wage incentives, or transparency of budget processes), as well as perception data on state credibility in the eyes of private and civic actors.

4.20 To gauge the impact of ID efforts on the *frontlines* properly, evaluators should report performance data on access to essential services, utilization, and client satisfaction. These data are usually collected through household surveys supported by the Bank in most client countries, including the Africa Region's Core Welfare Indicators Questionnaire (CWIQ). Another key frontline concern is operational efficiency, which comprises both unit costs as well as resource leakage. While such data were not always available in the past, several country teams have embarked on facility level surveys to gauge cost-efficiency. They have also started using expenditure tracking surveys to monitor the degree to which public resources are leaked or wasted before reaching facilities (Uganda).

4.21 Step 6: Judge efficacy. Once they have empirically established whether structural and capacity constraints were relieved (or institutional quality and frontline performance enhanced), evaluators should be able to judge *both intended and unintended consequences of ID efforts on the transformation process.* Depending on the quality of data generated by the scorecard, evaluators will also be able to estimate the opportunity costs—in terms of forgone improvements in public management systems—of adopting short-term, enclavist approaches to ID.

[65] Manning and Knack, 1999.

4.22 That said, judgments about efficacy are complicated by the multiple layers (for example, structural and capacity constraints, institutional quality, frontline performance) at which the interventions could register an impact. Table 4.2 elaborates on five "impact scenarios" that evaluators will likely confront in judging the efficacy of Bank support. For instance, ID efforts in scenario 2 relieve structural constraints and improve institutional quality; yet they fail to improve frontline delivery performance. Alternatively, scenario 4 represents the case where ID efforts fail to improve institutional quality, even through they correspond to gains in frontline performance The key is that a range of exogenous factors are at play during implementation of ID interventions. One approach to addressing such attribution problems is first to explicitly identify cases where the expected benefits of ID interventions in terms of frontline performance were likely undermined by exogenous factors. More important, CAEs should designate *improvements in institutional quality* (rather than frontline performance improvements) as the standard against which the efficacy of Bank-supported ID interventions is judged. In applying this standard to the five scenarios presented in Table 4.2, evaluators would deem Bank assistance effective only in scenarios 1 and 2.

4.23 Step 7. Attribute impact, identify downstream factors. Whether or not efforts are deemed effective, CAEs should be prepared to attribute impact (on structural factors, institutional quality, or operational efficiency) to the Bank's ability either "to do the right things" (that is, responsiveness and relevance) or to "do things right" (that is, leverage downstream factors during implementation). When ID efforts relieve structural and capacity constraints (for example, increased civil service pay scales for staff in economic management posts), but fail to improve either institutional quality (for example, policy volatility) or frontline performance, evaluators should be prepared to attribute the outcome to one of two causes.

4.24 First, the failure to improve institutional quality could mean that ID efforts targeted the "wrong" structural or capacity constraints. If so, evaluators should revisit their judgments on relevance and responsiveness. A second possibility is that Bank support was responsive to the right constraints, yet exogenous factors prevented interventions from influencing institutional

Table 4.1: Attributing Institutional Development Impact: Five Scenarios

If observe improvements in….				**Then….**	
Scenario	**Structural or capacity conditions**	**Institutional quality**	**Frontline performance**	**Indicates**	**Go to step**
1	Yes	Yes	Yes	ID efforts produced desired effect.	7
2	Yes	Yes	No	Which exogenous factors hindered improvements on the frontline?	7
3	Yes	No	No	Was Bank support relevant and responsive? Did ID efforts target the 'right' structural and capacity constraints?	2, 3
				Which exogenous factors prevented ID efforts from impacting institutional quality?	2, 3
				From impacting frontline performance?	7
4	Yes	No	Yes	Which exogenous factors hindered improvements on the frontline?	7
5	No	No	No	Was Bank support relevant and responsive? Did ID efforts target the 'right' structural and capacity constraints?	2, 3

quality. A third possibility could be that *downstream* factors such as moral hazard on the part of frontline workers prevented the potential benefits of well-designed interventions from registering at the facility level. This implies a need for improved design of implementation arrangements. The key is for evaluators to go through a systematic process of attributing the various levels of ID impact to downstream factors that are either endogenous or exogenous to institutional design.

4.25 Downstream factors that are typically significant in explaining the impact of ID effort include design of implementation arrangements, Bank supervision and monitoring, changing institutional ground realities, information asymmetries (and related moral hazard problems), political economy dynamics, and external shocks. The deleterious impact of complex implementation arrangements on project performance and ID impact, particularly in capacity-weak settings, has been identified in several OED evaluations of Bank-supported country and sector programs. An equally important *design* issue is the degree of flexibility afforded to clients during implementation. ID interventions in particular should be sufficiently flexible to enable clients to negotiate rapidly evolving institutional realities on the ground and to "learn by doing."

4.26 Typically, a *quid pro quo* for more flexible design is more reliable and timely monitoring of implementation progress. In addition to calculating supervision intensity on ID efforts with Bank-wide and regional averages, evaluators should closely examine the quality of supervision reports. The absence of timely *information* on the implementation of ID efforts offers greater scope for opportunism and moral hazard on the part of officials in borrower governments who may not be committed to the objectives of Bank-supported efforts. Evaluators should also consider whether supervision and monitoring functions were sufficiently decentralized for the Bank to be attuned rapidly changing realities on the ground.

4.27 *Political economy* factors such as those highlighted in Section 2 are likely to affect ID efforts as they are rolled out. Bounded rationality implies that political actors will continually re-assess their interests and shift alliances over the course of implementation. In other words, borrower ownership is not automatically “locked in” by formal, written agreements signed during negotiations with the Bank; it is dynamic. Finally, there is a growing need to model the role of external shocks in disrupting stable governance structures (as was the case in Albania and Indonesia). In such cases, CAEs should examine whether the Bank was prepared to handle the exigencies of crisis. For example, was there an early warning system or a standard operating procedure for country teams to alert senior management in the event of crises? There have been cases where internal processes such as high staff turnover significantly undermined the Bank’s ability to respond with agility to rapidly changing country conditions.

5. Summary and Next Steps

5.01 At the outset, this paper highlighted critical problems of scaling up, aggregation, and attribution that complicate country-level evaluations of ID impact. In attempting to address some of these concerns, the paper identified those institutions that “matter most” in terms of public sector transformation process; these comprise institutions governing policymaking, resource management, delivery, and regulation. It also summarized recent thinking on "good practice" approaches to structuring these public institutions and satisfying their capacity needs in order to maximize the allocative and technical efficiency of public expenditures.

5.02 Through this optic, the paper assessed a selection of recently completed CAEs and found urgent need for a standardized approach to benchmarking institutional performance, as well as more rigor in explaining the ID impact of Bank assistance. In addition to providing a detailed questionnaire and scorecard for rating structural and capacity constraints in the public sector, the paper offers a step-by-step approach for judging the responsiveness, relevance, and efficacy of Bank-supported ID at the country level.

5.03 Given its more pressing corporate objectives of "moving (evaluation) to a higher plane" and updating its approach to evaluating ID impact, OED should:

* Synchronize the proposed methodology of ID in this paper with ongoing work on developing an algorithm for country-level evaluations.

* Roll out an intensive training program for evaluators on country-level ID impact analysis once the methodology has been approved.

* Rapidly generate baseline data on country-level institutional performance, including structural and capacity conditions, institutional quality, and frontline performance. These data should provide the basis for more rigorous empirical analysis of ID approaches that are most likely to drive performance on the frontlines.

5.04 In developing training programs and the database, there is significant scope for collaboration with PREM, WBI, and Regional staff.

Bibliography

Africa Capacity Building Group. 1999. *Meeting with Ethiopian Leadership Group on Capacity Building*. Back-to-the-Office Report, Africa Capacity Building Unit (AFRTI2), Washington, D.C.: World Bank.

Alesina, Alberto, R. Hausmann, R. Hommes, and E. Stein. 1996. "Budget Institutions and Fiscal Performance in Latin America." NBER Working Paper No. 5556. National Bureau of Economic Research, Cambridge, Mass.

Aoki, Masahiko. 1998. "The Mechanisms of Institutional Change: Subjective Game-Form Framework of Analysis and Evidence from Japan." Closing Address, Second Annual Conference, Paris: International Society for New Institutional Economics.

Banks, Jeffrey, and Eric Hanushek. 1995. *Modern Political Economy: Old Topics, New Directions*. New York: Cambridge University Press.

Bennett, Sarah, Barbara McPake, and Anne Mills, eds. 1997. *Private Health Providers in Developing Countries*, London: Zed Books.

Bennett, Sarah, Steven Russell, and Anne Mills. 1996. "Institutional and Economic Perspectives on Government Capacity to Assume New Roles in the Health Sector: A Review of Experience." PHP Departmental Publication, Health Economics and Financing Programme, Health Policy Unit, London: London School of Hygiene and Tropical Medicine.

Bhatnagar, Bhuvan, and Aubrey C. Williams. *Participatory Development and the World Bank: Potential Directions for Change.* World Bank Discussion Paper No. 183, Washington, D.C.: World Bank.

Biddle, Jesse, and Vedat Milor. 1999. "Where Do Good Policies Come From?" Washington, D.C.: World Bank.

Burki, Shahid Javed, Guillermo Perry, and William Dillinger with Charles Griffin, Jeffrey Gutman, Fernando Rojas, Steven Webb, and Donald Winkler. 1999. *Beyond the Center: Decentralizing the State*. Washington, D.C.: World Bank.

Campos, Ed, and Sanjay Pradhan. 1996. *The Impact of Budgetary Institutions on Expenditure Outcomes*. Policy Research Working Paper, No. 1646, Washington, D.C.: World Bank.

Chhibber, Ajay. 1998. "Institutions, policies, and development outcomes," in *Evaluation and Development: The Institutional Dimension.* Eds. Robert Picciotto and Eduardo Wiesner. New Brunswick, N.J.: Transaction Publishers.

Clague, Christopher, Phillip Keefer, Stephen Knack, and Mancur Olson. 1996. "Property and Contract Rights under Democracy and Dictatorship" in *Journal of Economic Growth.* 1:2, 243-276 (June).

Dasgupta, Partha. 1999. "Poverty Reduction and Non-Market Institutions." Conference on Evaluation and Poverty Reduction, Washington, D.C.: World Bank.

Faguet, Jean-Paul. 1998. "Decentralization and Local Government Performance: Improving Public Provision in Bolivia." Center for Economic Performance and Development Studies Institute, London: The London School of Economics.

Girishankar, Navin. 1997. "A Review of Civil Service Reform in World Bank Economic and Sector Work." Background Paper, OEDCR, Washington, D.C.: World Bank.

_____. 1999. *Reforming Institutions for Service Delivery: A Framework for Development Assistance with an Application to the Health, Nutrition, and Population Sector.* Policy Research Working Paper No. 2039, January, Washington, D.C.: World Bank.

_____. 2000. "The Public Interest Under Pluralistic Institutional Design," in *Applying Public Administration to Development*, Paul Collins, ed. London: John Wiley & Sons, Ltd.

Girishankar, Navin, and Brian Levy. 2000. "Addressing Governance and Institutional Issues in the Poverty Reduction Strategy Process." AFTI2, Washington, D.C.: World Bank.

Girishankar, Navin, and Nick Manning. 1999. "Assessing Constraints on Frontline Delivery in Decentralized Settings." Draft Toolkit, AFR/PREM, Washington, D.C.: World Bank.

_____. 2000. "Sequencing in the Public Sector." PREM, Washington, D.C.: World Bank.

Goodin, Robert E., ed. 1996. *The Theory of Institutional Design.* Melbourne, Australia: Cambridge University Press.

Haggard, Stephan, and Mathew D. McCubbins, eds. 1997. *Political Institutions and the Determinants of Public Policy: When Do Institutions Matter?* San Diego, CA: University of California.

Hanushek, Eric, with Charles S. Benson, et al. 1994. *Making Schools Work: Improving Performance and Controlling Costs.* Washington, DC: Brookings Institute.

Johnson, Ronald N., and Gary Libecap. 1994. *The Federal Civil Service System and the Problem of Bureaucracy: The Economics and Politics of Institutional Change.* Chicago, IL.: University of Chicago Press.

Lawson, Andrew. 1999. "Public Expenditure Management: The Important Questions." Presentation, Oxford Policy Analysis, Redworth, UK.: DFID Governance Retreat.

Levy, Brian. 1999. "Patterns of Governance in Africa." Mimeo, Washington, D.C.: World Bank.

Lindauer, David, and Barbara Nunberg, eds. 1994. *Rehabilitating Government*. Regional and Sectoral Studies, Washington, D.C.: World Bank.

Lienert, Ian, and Jitendra Modi. 1997. "A Decade of Civil Service Reform in Sub-Saharan Africa," Working Paper, Washington, D.C.: International Monetary Fund.

Mackay, Keith. 1998. "The Development of Australia's Evaluation System." *Evaluation Capacity Development.* OED Working Paper Series: No. 4, May. Washington, D.C.: World Bank.

Manning, Nick, and Steve Knack. 1999. "Institutions Matter-But Which Institutions? Maintaining an Institutional Scorecard for Tracking Governance." PREM, September, Washington, DC: World Bank.

Manning, Nick, with Naazneen Barma, Jean Blondel, Elsa Pilichowski, and Vincent Wright. 1999. "Making the Cabinet Work: Institutional Arrangements for Strategic Decision-Making in Government." Administrative and Civil Service Reform Thematic Group Working Paper Series No. 1, May, Washington, D.C.: World Bank.

Manning, Nick, Naazneen Barma, Suzanne Dove, Nick Manning, Barbara Nunberg, Jana Orac, Gary Reid, and Jeffrey Rinne. 1999a. "Civil Service PREMpack." Draft Toolkit, December, Washington, D.C.: World Bank.

Mossier, Peter. 1996. "Checks and Balances, and the Supply of Central Bank Independence," Working Paper, University of St. Galleon, [location?].

Nunberg, Barbara, and John Nellis. 1990. *Civil Service Reform and the World Bank.* Discussion Paper No. 161, Washington, D.C.: World Bank.

Nunberg, Barbara. 1995. *Managing the Civil Service: Reform Lessons from Advanced Industrialized Countries.* Discussion Paper No. 204, Washington, D.C.: World Bank.

Operations Evaluation Department. 1997a. *1997 Annual Review of Development Effectiveness.* Washington, D.C.: World Bank.

_____. 1999a. *Civil Service Reform: A Review of World Bank Assistance.* Washington, D.C.: World Bank.

_____. 1998. *Evaluation Capacity Development in Africa: Selected Proceedings from a Seminar in Abidjan.* November. Washington, D.C.: African Development Bank and World Bank Operations Evaluation Departments.

_____. 1997b. "Evaluating Development Operations: Methods for Judging Outcomes and Impacts," *Lessons and Practices.* No. 11, November, Washington, D.C.: World Bank.

_____. 1999b. *The Impact of Public Expenditure Reviews: An Evaluation.* Washington, D.C.: World Bank.

Preker, Alexander, April Harding, and Navin Girishankar. 1999. "The Economics of Public and Private Participation in Health Care: New Insights from Institutional Economics." Washington, D.C.: World Bank. Processed.

Riker, William. 1982. "Implications from the Disequilibrium of Majority Rule for the Study of Institutions," in *Political Equilibrium.* Eds. Peter C. Ordeshook and Kenneth A. Shepsle. Boston: Kluwer-Nijhoff Publishing.

Shah, Anwar. 1998. "Balance, Accountability, and Responsiveness: Lessons about Decentralization." Policy Research Working Paper No. 2021, December, Washington, DC: World Bank.

Sabel, Charles. (1995). 'Design, Deliberation, Democracy: On the New Pragmatism of Firms and Public Institutions, Conference Presentation, Florence Italy: European University Institute.

Shepsle, Kenneth. 1998. "The Political Economy of State Reform: Political to the Core," Centennial Lecture, London: the London School of Economics.

Shepsle, Kenneth, and Weingast, B. (1981). "Structure-Induced Equilibrium and Legislative Choice," *Public Choice*, **37**, 503-519.

Simon, Herbert. 1983. *Reason in Human Affairs.* Stanford, Calif.: Stanford University Press.

Stapenhurst, Rick, and Ken Dye. 1998. “Pillars of Integrity: The Importance of Supreme Audit Institutions in Curbing Corruption,” in *New Perspectives on Combating Corruption.* Washington, DC: Transparency International and the World Bank Institute.

Stevens, Joe B. 1993. *The Economics of Collective Choice*. Boulder, Col.: Westview Press.

Williamson, Oliver E. 1985. *The Economics Institutions of Capitalism.* New York: The Free Press.

Wilson, James Q. 1990. *Bureaucracy: What Government Agencies Do and Why They Do It.* New York: Basic Books.

World Bank. 2000. *Reforming Public Institutions and Strengthening Governance.* A World Bank Strategy, Public Sector Group, Poverty Reduction and Economic Management (PREM) Network, November, Washington, D.C.: World Bank.

_____. 1997. *World Development Report 1997: The State in a Changing World.* New York: Oxford University Press.

_____. 1999. *World Development Report 1999: Localization and Globalization.* New York: Oxford University Press.

Annex 1: Country Assistance Evaluations Reviewed

Table 1-A: Sample of CAEs

Region	CAE	Task Manager	Date to CODE
AFR	Burkina Faso	A. Galenson	November 19, 1999
	Ethiopia	A. Ray	June 17, 1999
	Kenya	P. Gupta	May 14, 1998
EAP	Indonesia	R. Vandendries	February 5, 1999
	Thailand	L. Goreux	June 17, 1998
ECA	Albania	R. Robinson	June 18, 1998
MNA	Yemen	L. Effron	March 30, 1999
LAC	Bolivia	L. Ramirez	June 5, 1998
	Jamaica	A. Galenson	December 30, 1998
SAS	Nepal	A. De Silva	December 3, 1998

Table 1-B: Matrix for assessing CAEs

Criteria	Policy	Resources	Delivery	Regulation
COVERAGE 1. Institutions as integral or incidental 2. Which institutions matter?				
RELEVANCE a/ Initial diagnoses: qualitative aspects of performance that matter? b/ Goodness of fit to existing capacity c/ Political economy and points of entry d/ Design preserves relevance? e/ Metrics?				
EFFICACY a/ A metric b/ Attribution -- projects affecting sectoral performance -- did the Bank make a difference? What's the counterfactual? -- factors endogenous or exogenous to Bank/situation				
RECOMMENDATIONS a/ Follow through logically b/ Implicit or explicit impact on ID				
CROSS-CUTTING THEMES a/ Why do institutions change—technocratic knowledge or politics, both? b/ CSR as a proxy for institutional reform c/ Other comments				

Annex 2. Mapping Constraints and Interventions

1. This annex provides a detailed questionnaire for evaluators to rate the severity of structural and capacity constraints on policymaking, resource management, delivery, and regulatory enforcement using a simple six-point scale (0-5). It also enables evaluators to assess the intensity of Bank-supported structural reform and capacity building efforts. By superimposing a map of interventions onto a map of initial constraints, CAEs should be able to judge more systematically the responsiveness of Bank assistance.

Diagnosing Initial Institutional Constraints

2. Each of the questions provided below has been coded to indicate structural (S) and capacity (C) constraints, as well as their impact on institutional quality (Q). It is envisaged that data generated by questionnaire and accompanying scorecard will be used to assist in ongoing CAEs and serve as benchmarks for future evaluations.

Credible Policymaking and Planning

3. *Macroeconomic* (monetary and fiscal) policies are most likely to be predictable, consistent, and coherent when policymaking processes are insulated. Typically, technocratic enclaves with arms-length relationships to the executive serve to buffer policymakers from populist pressures such as interest group capture. The following questions enable evaluators to rate the severity of initial constraints on macroeconomic policymaking.

i) To what extent was macroeconomic policymaking exposed to political interference by the executive or pressures from other interests? (S)

0	**1**	**2**	**3**	**4**	**5**
Highly insulated	Insulated	Somewhat insulated	Somewhat exposed	Exposed	Highly exposed

ii) Were there any formal mechanisms to insulate monetary and fiscal policy from such pressures? (S) Yes ___ No ___

iii) If yes, which ones (mark an F for fiscal and an M for monetary)? (S)
Delegation to an autonomous entity _____ Financially endowed enclaves ____
Fixed policy and bargaining rules _____ Binding international agreements _____

iv) How frequently did politicians, bureaucrats, and other stakeholders abrogate formal rules and due process during monetary and fiscal policymaking? (S)

0	**1**	**2**	**3**	**4**	**5**
Never	Rarely	Occasionally	Frequently	Almost always	Always

v) In your judgment, did structural constraints affect macroeconomic policy outcomes in any of the following ways? Score on 0-5 scale, where 0=no impact and 5=significantly increased. (S)
Volatility _____ Inconsistency _____ Incoherence_____

vi) Estimate vacancy rate for managerial and technical posts in the central bank, Ministry of Finance, and other key technocratic enclaves. (C)
>50% _____ 35-50% _____ 20-35%_____ 5-15%_____ <5% _____

vii) To what extent were professional staff-in-post adequately skilled to carry out core economic management tasks? (C)

0	1	2	3	4	5
Never under-skilled	Rarely under-skilled	Occasionally under-skilled	Frequently under-skilled	Almost always under-skilled	Chronically under-skilled

viii) What was the quality of the statistical base available to policymakers to carrying out core economic management tasks? (C)

0	1	2	3	4	5
Very robust	Robust	Somewhat robust	Somewhat weak	Weak	Very weak

ix) To what extent were day-to-day functions of core policymaking bodies constrained by scarcity of complementary inputs such as hardware and software, infrastructure, stationery, etc.? (C)

0	1	2	3	4	5
Completely unconstrained	Unconstrained	Somewhat unconstrained	Somewhat constrained	Constrained	Severely constrained

x) In your judgment, did capacity constraints affect macroeconomic policy outcomes in any of the following ways? Score on 0-5 scale, where 0=no impact, and 5=significantly increased. (Q)
Volatility _____ Inconsistency _____ Incoherence_____

4. The credibility of *sectoral* policies is enhanced through inclusive—rather than insular—processes. Open and inclusive policymaking enables multiple stakeholders to reveal their preferences and debate alternatives. The questions that follow help evaluators rate the severity of structural and capacity constraints on sectoral policymaking processes.

i) To what extent were sectoral policymaking processes closed to participation by a broad range of stakeholders? (S)

0	1	2	3	4	5
Completely open	Open	Somewhat open	Somewhat closed	Closed	Completely closed

ii) Which of the following mechanisms, if any, were in place to involve multiple stakeholders in sectoral policymaking? (If applicable, circle one or more of the options on the right) (S)

Sector working groups _____	If yes>>	Inter-ministerial	Local gov't	Donors
Sub-national committees _____	If yes>>	Provincial	District	Sub-district
Deliberation council _____	If yes>>	Private	Public	Non-profit
Other _____				

iii) In which of the following aspects of sectoral policymaking were sub-national or civil society actors with longer-term interests typically involved? (S)
System-wide goals _____ Content standards _____
Performance standards and benchmarks ____ Types of services to be provided_____
Role of private and civic actors_____ Other_____ None of the above _____

iv) To what extent were these policy and planning activities captured (by special interests or because of exclusionary practices)? (S)

0	1	2	3	4	5
Completely open and transparent	Open and transparent	Somewhat open and transparent	Somewhat captured	Captured and closed	Completely captured and closed

v) In your judgment, did closed or captured planning processes affect sectoral policy outcomes in any of the following ways? Score on 0-5 scale, where 0=not at all and 5=very severely. (Q)
Volatility _____ Inconsistency _____ Incoherence_____

vi) Estimate the vacancy rate for managerial and technical posts in line agencies. (C)
>50% _____ 35-50% _____ 20-35%_____ 5-15%_____ <5% _____

vii) To what extent were professional staff-in-post under-skilled to carry out basic line agency functions? (C)

0	**1**	**2**	**3**	**4**	**5**
Never under-skilled	Rarely under-skilled	Occasionally under-skilled	Frequently under-skilled	Almost always under-skilled	Chronically under-skilled

viii) What was the quality of the statistical base available to carry out sectoral policymaking and planning? (C)

0	**1**	**2**	**3**	**4**	**5**
Very robust	Robust	Somewhat robust	Somewhat weak	Weak	Very weak

ix) To what extent were the day-to-day functions of line agencies constrained by scarcity of complementary inputs such as hardware and software, infrastructure, stationery, etc.? (C)

0	**1**	**2**	**3**	**4**	**5**
Completely unconstrained	Unconstrained	Somewhat unconstrained	Somewhat constrained	Constrained	Severely constrained

x) How severely did each of the following capacity constraints hinder sub-national authorities from participating in policymaking? Score on 0-5 scale, where 0=not at all and 5=severely hindered. (C)
Poor knowledge of sector policy __ Unfamiliarity with procedures _____
Absence of sector specialists at the sub-national level (for example, water engineers) ____
Lack of information on sectoral performance _____
Lack of basic literacy _____

xi) How severely did each of the following capacity constraints hinder communities from participating in planning? Score on 0-5 scale, where 0=not at all and 5=severely hindered. (C)
Poor knowledge of their entitlements in given sector _____ Lack of basic literacy _____
Unfamiliarity with planning bodies _____ Informal exclusionary practices _____

xii) In your judgment, did capacity constraints impact sectoral policy outcomes in any of the following ways? Score on 0-5 scale, where 0=no impact and 5= significantly increased. (Q)
Volatility _____ Inconsistency _____ Incoherence_____

xiii) Make an overall judgment of the impact of structural constraints on policy credibility.

0	**1**	**2**	**3**	**4**	**5**
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

xiv) Make an overall judgment of the impact of capacity constraints on policy credibility.

0	**1**	**2**	**3**	**4**	**5**
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

Adequate and Predictable Resource Management

5. Transparent yet hierarchical processes promote adequate, predictable *financial resource flows* to implementing agencies. They ensure that key decisionmakers do not renege on agreements over inter- and intra-sectoral allocations during budget execution. They also limit uncertainty in budget and personnel management. CAEs should rate the degree of which budgetary processes meet these conditions.

i) To what extent was budget execution governed by transparent and hierarchical processes, both of which have been associated with favorable fiscal performance? (S)

	0	**1**	**2**	**3**	**4**	**5**
Degree of transparency	Highly transparent	Generally transparent	More transparent than opaque	More opaque than transparent	Opaque	Highly opaque
	0	**1**	**2**	**3**	**4**	**5**
Degree of hierarchy	Sufficiently hierarchical	Generally hierarchical	More hierarchical than collegial	More collegial than hierarchical	Collegial	Highly collegial

ii) How likely was it for cabinet-level actors to renege on agreements regarding inter-sectoral resource allocations in the middle of the budget cycle? (S)

0	**1**	**2**	**3**	**4**	**5**
Very unlikely	Unlikely	Somewhat unlikely	Somewhat likely	Likely	Very likely

iii) When reallocations are made during the budget cycle, where are they most likely to be cut first? Circle one or more. (S)
Social sectors _____ Infrastructure _____ SOEs _____ Economic management _____
Defense and police _____ Other _____

iv) If there were greater certainty that budgets would be executed as planned, would original allocations for any of the following items increase? Circle one. (S)
Wage _____ Non-wage recurrent _____ Capital _____

v) How likely was it that resources would be leaked during budget execution? (S)

0	**1**	**2**	**3**	**4**	**5**
Very unlikely	Unlikely	Somewhat unlikely	Somewhat likely	Likely	Very likely

vi) How significant were each of the following factors in contributing to resource leakage? Score on 0-5 scale, where 0=not at all significant and 5=very significant. (S)
Unclear rules procedures for public financial accounting _____
Lack of capacity to account for budgeting resources _____
Weak internal and external auditing practices _____
Weak sanctions against corrupt practices _____

vii) At which level of public management were resources most likely to be leaked? (S)
Between MoF and line ministries _____ Within line ministries_____
Within sub-national administrative units____ At the facility level _____

viii) To what extent were off-budget sources of funding used to compensate for budgetary uncertainty? (S)

Sectors	Funding source (private/corporate, NGO, users, etc)	% of capital	% of recurrent
(1)			
(2)			

ix) Were any of the following mechanisms used to publicize financial accounts and audits? (S)
Compliance with Freedom of Information Act _____
Regular internal auditing, sent to MoF _____
Independent audits subject to legislative oversight _____

x) In your judgment, did structural constraints affect budgetary flows in any of the following ways? Score on 0-5 scale, where 0=not at all and 5=very severely. (Q)
Variance (between actual and budgeted funds) _____
Expectations of variance _____
Leakage _____

xi) Estimate the vacancy rate for posts directly involved in budgeting and accounting. (C)
>50% _____ 35-50% _____ 20-35%_____ 5-15%_____ <5% _____

xii) To what extent were professional staff-in-post under-skilled to carry out basic financial management tasks? (C)

0	**1**	**2**	**3**	**4**	**5**
Never under-skilled	Rarely under-skilled	Occasionally under-skilled	Frequently under-skilled	Almost always under-skilled	Chronically under-skilled

xii) In your judgment, did capacity constraints affect budgetary flows in any of the following ways? Score on 0-5 scale, where 0=not at all and 5=very severely. (Q)
Volatility _____ Inconsistency _____ Incoherence_____

xiii) Did inadequate and unpredictable budgetary flows in turn contribute to any of the following delivery problems? Score on 0-5 scale, where 0=not at all and 5=very severely. (Q)
Problems in hiring and keeping qualified personnel _____
Wage arrears_____
Problems in meeting demand for services _____
Arrears to procurement suppliers _____
Cancellation of procurement contracts _____
Problems in securing procurement contracts _____

6. Equally important are the availability and proper deployment of *human resources* in key government institutions. Attracting and retaining qualified personnel requires that public institutions offer incentives that are competitive, as well as generally favorable conditions of service. These include fostering a work environment conducive to professional development and insulated from political interference or patronage.

i) Estimate the vacancy rates that affected managerial and technical posts in the following institutions? Circle one. (S)

Core economic management ministries	<5%	5-15%	15-20%	20-35%	35-50%	>50%
Sectoral ministries	<5%	5-15%	15-20%	20-35%	35-50%	>50%
Sub-national institutions (highest tiers)	<5%	5-15%	15-20%	20-35%	35-50%	>50%
Sub-national institutions (lower tiers)	<5%	5-15%	15-20%	20-35%	35-50%	>50%

ii) Was overstaffing evident among the semi-skilled and low-skilled job classes? (S)
Yes___ No___

iii) What were the approximate wage decompression ratio between the highest and lowest grades, between managerial and technical staff? (S)

Highest-to-lowest	<5:1	5:1	7:1	10:1	15:1	20:1	25:1	30:1	35:1	>35:1
Managerial-to-technical	<5:1	5:1	7:1	10:1	15:1	20:1	25:1	30:1	35:1	>35:1

iv) How did the wage of public sector employees compare as a percentage of those of private sector comparators? Indicate relative wages for the different skill levels listed below. (S)

Managerial	<20%	20-35%	35-50%	50-65%	65-80%	80-95%	>95%
Technical	<20%	20-35%	35-50%	50-65%	65-80%	80-95%	>95%
Semi-skilled	<20%	20-35%	35-50%	50-65%	65-80%	80-95%	>95%
Low-skilled	<20%	20-35%	35-50%	50-65%	65-80%	80-95%	>95%

v) To what extent were grading systems too complex, rigid, and therefore opaque? (S)

0	**1**	**2**	**3**	**4**	**5**
Sufficiently streamlined	Generally streamlined	Somewhat streamlined	Somewhat complex	Complex	Highly complex

vi) To what extent were line agencies and key civil service management institutions such as the Ministry of Public Service or Public Service Commission understaffed? (C)

0	**1**	**2**	**3**	**4**	**5**
Never	Rarely	Occasionally	Frequently	Almost always	Always

vii) What if any impact did the HIV/AIDS pandemic have on medium-term prospects for retaining and developing capacity in the civil service? (C)

0	**1**	**2**	**3**	**4**	**5**
Not at all	Insignificant	Mild	Moderate	Significant	Very significant

viii) What if any impact did pay and systems have on vacancy rates in the civil service? (Q)

0	**1**	**2**	**3**	**4**	**5**
Not at all	Insignificant	Mild	Moderate	Significant	Very significant

ix) To what extent did pay and grading systems contribute to the following problems? Score on 0-5 scale, where 0=not at all and 5=very significant. (Q)
Grade creep _____
Patronage hiring in lower posts_____
Ghosts _____
Other _____

x) In your judgment, did difficulties in retaining and attracting qualified staff hinder government agencies from carrying out core functions such as budgeting, service delivery, and M&E? (Q)

0	**1**	**2**	**3**	**4**	**5**
Not at all	Insignificant	Mild	Moderate	Significant	Very significant

xi) Make an overall judgment of the severity of structural constraints on the adequacy and predictability of human and financial resource flows.

0	1	2	3	4	5
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

xii) Make an overall judgment of the severity of capacity constraints on the adequacy and predictability of human and financial resource flows.

0	1	2	3	4	5
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

Flexible Delivery

7. Delivery systems that are flexible are better able to produce outputs efficiently. Flexible delivery systems typically decentralize accountability for results (for example, for the provision of health services) along with the authority to control fiscal and administrative capacity. Once decentralized, service provision should also be de-monopolized. In other words, providers should also be subjected to a richer combination of checks and balances such as competition and beneficiary participation. CAEs should assess both the structural flexibility in delivery systems as well as the capacity of relevant public institutions to manage increased institutional pluralism.

i) Were any responsibilities for service delivery assigned to sub-national tiers of government? (S) Yes___ No ___

ii) How much scope was there for further decentralization of service delivery, particularly of local public goods, to lower levels of government? (S)

0	1	2	3	4	5
None at all	Very little	Little	Some	Significant	A great deal

iii) If certain delivery responsibilities were decentralized, how much authority or discretion did sub-national units have to allocate financial and human resources or other inputs? (S)

0	1	2	3	4	5
A great deal	Significant	Some	Little	Very little	None at all

iv) To what extent where sub-national authorities constrained in exercising discretion over the following aspects of pay policy and personnel management? Score on a 0-5 scale, where 0=not at all constrained and 5=highly constrained. (S)
Hiring and firing personnel_____ Career development (seniority vs. internal competition) ____
Term of employment (individually-determined contracts vs. uniform) _____
Setting compensation level (job evaluation vs. market-based) _____

v) Did central authorities institute a system of conditional grants to sub-national government in order to ensure delivery of essential services? (S) Yes___ No___

vi) If yes, how severely did the conditionality in intergovernmental transfers constrain sub-national authorities from making legitimate allocative decisions? Score on a 0-5 scale, where 0=not at all constrained and 5=highly constrained. (S)
Between wage and non-wage recurrent_____ Between recurrent and capital _____
Between items within non-wage recurrent_____

vii) Were sub-national governments legally authorized to raise own-source revenues (including user fees) to supplement transfers from the Consolidated Fund? (S)

0	1	2	3	4	5
A great deal	Significant	Some	Little	Very little	Not at all

viii) Did the regulatory framework provide for service delivery under any of the following institutional options? If yes, designate the levels of government for which the law applies. (S)

Private sector participation _____	Federal	Provincial	District
Contracting out or outsourcing _____	Federal	Provincial	District
Autonomous public sector bodies _____	Federal	Provincial	District
NGO participation _____	Federal	Provincial	District
Turnover to users association _____	Federal	Provincial	District

ix) In your judgment, did these institutional constraints on flexibility adversely affect frontline performance in any of the following ways? Score on 0-5 scale, where 0=no impact and 5=severely undermined. (Q)

Usage _____ Cost-efficiency_____ Timeliness _____ Client satisfaction_____

x) Rate the severity of the following capacity constraints on public institutions at the sub-national level. Score on a 0-5 scale, where 0=not at all; 5=severely constrained. (C)

Personnel in: Financial mgt._____ Procurement ____ Audit _____ Regulation ____ Frontline _____
Statistical base for: Fin. mgt._____ Pay and personnel_____ Frontline performance_____
Complementary inputs: IT _____ Infrastructure _____ Vehicles_____ Other _____

xi) Rate the severity of the following capacity constraints on private and civic actors to undertake service delivery activities. Score on a 0-5 scale, where 0=not at all; 5=severely constrained. (C)

Personnel in: Financial mgt.___ Procurement _____ Frontline (e.g., teachers) _____
Statistical base for: Fin. mgt._____ Pay and personnel_____ Frontline performance_____
Complementary inputs: IT _____ Infrastructure _____ Vehicles_____ Other _____

xii) In your judgment, did these capacity constraints on flexibility adversely affect frontline performance in any of the following ways? Score on 0-5 scale, where 0=no impact and 5=severely undermined. (Q)

Utilization _____Cost-efficiency_____ Timeliness _____ Client satisfaction_____

xiii) Make an overall judgment of the severity of structural constraints on flexibility in delivery systems.

0	1	2	3	4	5
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

xiv) Make an overall judgment of the severity of capacity constraints on flexibility in delivery systems.

0	1	2	3	4	5
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

Enforceable Regulation

8. Enforceable "regulation"—namely, controls over inputs, benchmarks for outputs, and evaluation of outcomes—not only safeguards the transformation process from various forms of arbitrariness, but also promotes "learning by doing" to support continual improvement. It is critical that CAEs assess the appropriateness of regulation as well as credible commitment of key economic actors to these formal rules governing public management functions.

i) How would you characterize input controls in line agencies, i.e., controls over financial and human resources as well as physical assets? (S)

0	**1**	**2**	**3**	**4**	**5**
Very tight	Tight	Somewhat tight	Somewhat loose	Loose	Very loose

ii) Which of the following bodies were formally mandated with undertaking financial and operational audits? (S)
Supreme audit institutions (Auditor General or equivalent)____ Inspector General_____
Internal audit bodies _____ Non-governmental watchdog groups _____ Other _____

iii) Were budgetary institutions governed by any of the following input controls? (S)
Organic finance act _____ Treasury instructions _____
Budget circulars _____ Internal audits _____
External audit _____ Anti-Corruption Bureau or equivalent _____
Regular expenditure tracking surveys_____
Codes of ethics with specific provisions for fiduciary responsibility____
Reporting by external audit authority to parliamentary sub-committee _____
Publication of audit findings in the media _____

iv) Was the independence of external audit or M&E bodies weakened in any of the following ways? Score on 0-5 scale, where 0=not at all and 5=severely weakened. (S)
Unclear legal mandate for the institution _____
Weak legal provisions for access to information _____
Inadequate pay for staff _____
Compromised by source of funding _____
Aggregate resource scarcity _____

v) How delayed were budgetary institutions in meeting their fiduciary responsibilities in the following areas? Circle one for each item. (S)

In-year accounting	<6 months	9 months	12 months	18 months	24 months	36 months
Internal audits	<6 months	9 months	12 months	18 months	24 months	36 months
External audits	<6 months	9 months	12 months	18 months	24 months	36 months

vi) In the last 3-5 years, how typical was it for the legislature to examine formally and act on annual reports from the external audit authority (for example, the Auditor General)? (S)

0	**1**	**2**	**3**	**4**	**5**
Very typical	Typical	Somewhat typical	Somewhat atypical	Atypical	Very atypical

vii) Which of the following were used to monitor frontline performance in the last 3-5 years? (S)
Household surveys of client satisfaction _____ Report cards at the facility level _____
Firm-level surveys of service quality _____ Facility-level surveys of cost-efficiency ___
Village-level or local-level participatory evaluation _____ Other _____

viii) Were the data gathered from these exercises used in any of the following ways? (S)
Dissemination of data to officials across districts_____
Publication in local and national media _____
Incorporation in contractual agreements with heads of departments _____
Other _____

ix) To what extent was M&E given due importance as a management tool across government agencies and departments? (S)

0	**1**	**2**	**3**	**4**	**5**
Given high priority	Given priority	Given some priority	Somewhat ignored	Ignored	Completely ignored

x) Was it adequately linked into the policymaking and budgeting process? (S) Yes ____ No _____

xi) Using the three criteria listed below, rate the degree to which the independence of audit and evaluation bodies was routinely compromised. (S)

Arms-length status from government	**0**	**1**	**2**	**3**	**4**	**5**
	Not at all	Very little	Little	Some	Significant	A great deal
Access to information	**0**	**1**	**2**	**3**	**4**	**5**
	Not at all	Very little	Little	Some	Significant	A great deal
Sources of funding	**0**	**1**	**2**	**3**	**4**	**5**
	Not at all	Very little	Little	Some	Significant	A great deal

xii) In your judgment, did these capacity constraints weaken any of the elements of intra-public sector regulation? Score on 0-5 scale, where 0=not at all and 5=severely weakened. (Q)
Checks and balances on arbitrariness (input controls) _____
"Learning by doing" or continuous improvement (output benchmarks) _____
Understanding outcomes and feedback into policymaking and budgeting _____

xiii) Estimate the vacancy rate for posts in supreme audit institutions. (C)
>50% _____ 35-50% _____ 20-35%_____ 5-15%_____ <5% _____

xiv) Estimate the vacancy rate for accounting posts in line agencies including internal audit. (C)
>50% _____ 35-50% _____ 20-35%_____ 5-15%_____ <5% _____

xv) Estimate the vacancy rate for evaluation posts in line agencies. (C)
>50% _____ 35-50% _____ 20-35%_____ 5-15%_____ <5% _____

xvi) To what extent were professional staff-in-post under-skilled to carry out basic accounting and internal auditing? (C)

0	**1**	**2**	**3**	**4**	**5**
Never under-skilled	Rarely under-skilled	Occasionally under-skilled	Frequently under-skilled	Almost always under-skilled	Chronically under-skilled

xvii) To what extent were professional staff-in-post in SAIs or watchdog institutions adequately skilled to carry out basic public sector financial auditing functions? (C)

0	**1**	**2**	**3**	**4**	**5**
Never under-skilled	Rarely under-skilled	Occasionally under-skilled	Frequently under-skilled	Almost always under-skilled	Chronically under-skilled

xviii) Were key budgetary institutions equipped with a functioning IFMIS? (C) Yes ____ No ___

xix) If yes, which agencies or departments were networked? (C)
Core economic management bodies _____ Sectoral ministries _____
Sub-national government to the district level _____ Sub-district level _____

xx) Were "live" databases with data on sectoral output performance collected in the course of routine M&E? (C) Yes _____ No_____

xxi) Rate the level of capacity available to maintain government information systems. (C)

0	**1**	**2**	**3**	**4**	**5**
Nonexistent	Insignificant	Modest	Moderate	Significant	Very significant

xxii) How willing were local beneficiaries to be trained to undertake participatory evaluation in key sectors? (C)

0	**1**	**2**	**3**	**4**	**5**
Very willing	Willing	Somewhat willing	Somewhat reluctant	Reluctant	Very reluctant

xxiii) How much scope did professional evaluators in the private and public sectors have for career development? (C)

0	**1**	**2**	**3**	**4**	**5**
Nonexistent	Insignificant	Modest	Moderate	Significant	Very significant

xxiv) Assess the "skills gap" in the following discrete areas of intra-public sector regulation. Score on a 0-5 scale, where 0=no skills gap and 5=severe skills gap. (C)
Financial auditing _____ Operational or performance auditing _____ Benchmarking _____
Survey design _____ Participatory evaluation methods _____ Impact evaluation _____

xxv) In your judgment, did these capacity constraints weaken any of the elements of intra-public sector regulation? Score on 0-5 scale, where 0=not at all and 5=severely weakened. (Q)
Checks and balances on arbitrariness (input controls) _____
"Learning by doing" or continual improvement (output benchmarks) _____
Understanding outcomes and feedback into policymaking and budgeting _____

xxvi) Make an overall judgment of the severity of structural constraints on the enforceability of regulation.

0	**1**	**2**	**3**	**4**	**5**
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

xxvii) Make an overall judgment of the severity of capacity constraints on the enforceability of regulation.

0	**1**	**2**	**3**	**4**	**5**
Nonexistent	Insignificant	Mild	Moderate	Severe	Very severe

Mapping Constraints and Interventions

9. Using the scoring system suggested above, evaluators can calculate the "average severity" of structural and capacity constraints along the four axes of policymaking, resource management, delivery, and regulation. By way of example, Table 2-A illustrates how this approach can pinpoint those aspects of the transformation process that are weakest.[1] The table maps a public management system that is structurally constrained in two areas—resource management and regulatory enforcement. Specifically, these constraints comprise inadequate incentives, opaque and collegial budget processes, and weak restraints on the use of inputs. At the same time, the system is capacity-constrained in policymaking and service delivery. Ideally, the reform agenda in this setting should involve restructuring of institutional arrangements that govern budget formulation and execution, as well as capacity building for policymaking bodies and service providers. Section 4 provides further elaboration on how this type of institutional map can help chart a critical path for ID that is based on an appropriate sequence of structural reform and capacity building interventions.

Table 2-A: An Example of Scoring the Severity of Initial Constraints

(1) CONSTRAINTS ON TRANSFORMATION	(2) STRUCTURAL	(3) CAPACITY
CREDIBILITY OF POLICYMAKING		
Macroeconomic policymaking	3.00	3.00
Sectoral policymaking	2.00	4.00
Ave. severity of constraints	**2.50**	**3.50**
PREDICTABILITY OF RESOURCES		
Budgeting and fin. mgt.	4.00	2.00
Personnel	5.00	3.00
Ave. severity of constraints	**4.50**	**2.50**
FLEXIBILITY IN DELIVERY		
Decentralization	1.00	4.00
De-monopolization	4.00	3.00
Ave. severity of constraints	**250**	**3.50**
ENFORCEABILITY OF REGULATION		
Input controls	5.00	3.00
Output benchmarks	2.50	5.00
M&E to understand outcomes	2.50	1.00
Ave. severity of constraints	**3.33**	**3.00**

10. Mapping constraints is only the first step. A second step is for evaluators to categorize Bank-supported interventions, score their intensity, and then superimpose these on the "constraints map" in order to evaluate *responsiveness*. The intensity of interventions should also be scored on a 0-5 scale, along the same axes. Table 2-B describes a situation where Bank support focused primarily on building capacity (that is , the supply side) for policymaking, resource management, and to some extent, regulation. Support for structural reform—especially in the priority areas indicated by the constraints map—was insufficient. By superimposing Table 2-B onto 2-A, evaluators would conclude that the Bank was only marginally responsive. In other words, Bank assistance overemphasized capacity concerns when critical constraints were in fact structural. Even though policymaking and resource management did benefit from capacity building efforts, delivery institutions were largely ignored.

[1] A higher score indicates more severe constraints on those elements of transformation.

Table 2-B: An Example of Categorizing and Scoring the Intensity of Bank Support

(1) CORE ELEMENTS OF TRANSFORMATION	(2) STRUCTURAL	(3) CAPACITY	(4) BANK INSTRUMENT
CREDIBILITY OF POLICYMAKING			
Macroeconomic policymaking	1.00	4.50	SAL, TAL
Sectoral policymaking	2.50	3.00	SECAL
Ave. severity of constraints	**1.75**	**3.75**	
PREDICTABILITY OF RESOURCE FLOWS			
Budgeting and fin. mgt.	2.00	2.00	TAL
Personnel	2.00	4.50	SAL
Ave. severity of constraints	**2.00**	**3.00**	
FLEXIBILITY IN DELIVERY			
Decentralization	1.00	--	SECAL, TAL
De-monopolization	**--**	--	
Ave. severity of constraints	**0.50**	**0.00**	
ENFORCEABILITY OF REGULATION			
Input controls	3.00	4.00	TAL
Output benchmarks	--	2.00	LIL
M&E to understand outcomes	1.00	2.00	SAL, APL, LIL
Ave. severity of constraints	**1.33**	**2.67**	

Calculating Response Rates

11. This assessment of the Bank's responsiveness can be quantified in the form of *response rates.* One approach, illustrated in Table 2-C, is to take the simple ratio of average intensity of interventions to the average severity of constraints. It provides an indicator *inter alia* of the Bank's capacity to go beyond broad statements of strategic intent and actually design relevant interventions.[2]

Table 2-C: Response Rates

	STRUCTURAL	CAPACITY
Policy	70%	107%
Resources	44%	120%
Delivery	20%	0%
Regulation	40%	89%

[2] OED's 1999 Civil Service Reform Study provides an example—albeit a cruder one—of calculating response rates as the ratio of appropriate interventions designed to diagnoses of bureaucratic dysfunction.