Part II
Governance as Part of Global Monitoring

The emerging global architecture to support poverty reduction rests on the principle of mutual accountability. Donor countries are to be accountable for providing aid in ways that support country development strategies. Developing countries are to be accountable for using aid and other resources effectively. But, as this Global Monitoring Report (GMR) will detail, accountability with regard to aid resources is only a small part of the governance agenda. Both donor and developing countries are to be accountable more broadly for enhancing the checks and balances fundamental for development.

Part II of this GMR will spotlight these governance facets of the new architecture. The objective is to consider how to approach a satisfactory monitoring framework in the field of governance that is relevant for the scaling up of aid. This requires some analysis of relevant aspects of governance, a review of available data, and the use of some simple typology to allow for the wide differences in country situations.

There is strong evidence of a link between the quality of a country's governance system and its development performance: Empirical evidence links governance to growth, over time—

as, for example, in the thousand-year evolution of governance systems that underpin today's developed countries—and across countries, though some debate persists as to causality.

Statistical evidence suggests that the causality between growth and governance is two-way—implying that gains in either can give momentum to a virtuous spiral of development improvement. Figure II.1 provides evidence of the causality from governance to growth: it illustrates the statistically robust partial relationship (controlling for initial income and schooling levels) between the quality of governance across developing countries in 1982 and income growth over the subsequent two decades.[1]

Econometric studies show that the benefits of public health spending on child and infant mortality rates are greater in countries with better governance—and that, as countries improve their governance, public spending on primary education becomes more effective in increasing primary education attainment (Swaroop and Rajkumar 2002).

The scale of corruption has also posed extraordinary costs on some countries. A conservative estimate is that the former president of Zaire looted the treasury of some

FIGURE II.1 Governance and growth, 1982–2002

Per capita income growth, 1982–2002 (residual)

Initial governance, 1982
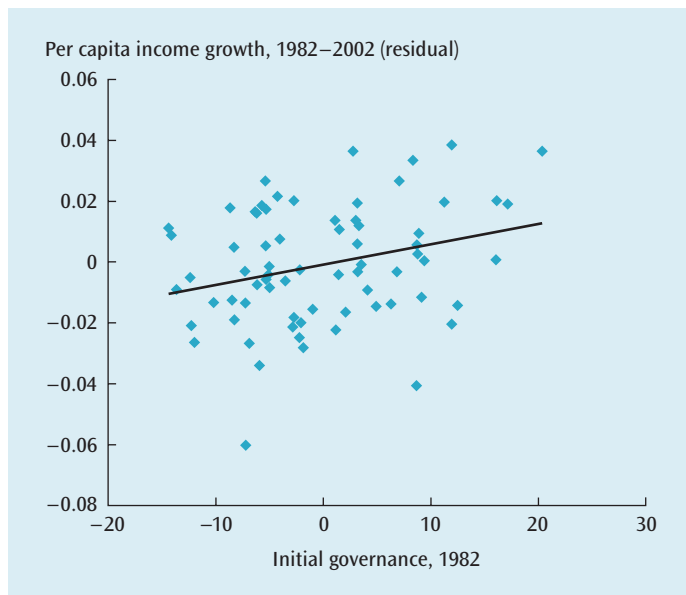
Source: Steve Knack, 2005.
Note: Governance measure is an index from the International Country Risk guide (ICRG). The growth estimates are the unexplained residuals after controlling for the impact of initial levels of income and education.

US$5 billion—an amount equal to the country's entire external debt at the time he was ousted in 1997. The funds allegedly embezzled by former presidents of Indonesia and Philippines are estimated to be two and seven times as high, respectively.[2] Micro-level studies reveal the ubiquitous daily impact of corruption—and the benefits of scaling it back. In health care, for example, during the first nine months of a 1996–7 crackdown on corruption in Buenos Aires, Argentina, the prices paid for basic inputs at public hospitals fell by 15 percent. In customs, the use of private international firms to conduct preshipment inspection of imports has been associated with increases in the growth rate of import duties of 6 to 8 points annually.[3]

Many elements of the "development checklist" are governance related. Are there mechanisms in place to ensure that public resources reach their intended purpose with little leakage? Is the investment climate supportive of growth and reductions in income poverty? Can countries develop plans and do they have the institutional capacity to execute them? Is there adequate information and transparency in government to foster the active civil society to build greater accountability? Are the incentives and accountabilities of teachers and health care workers adequate to ensure low absenteeism and shirking? Does the rule of law protect the rights of citizens? The answer to these questions depends on the quality of national governance systems.

Getting governance of a quality needed to meet the Millennium Development Goals (MDGs) is not simply a matter for aid recipients. The global milieu has powerful influences on the governance system in developing countries. Global markets can be the source of virulent, corrosive corruption—or a powerful disciplining device, helping to strengthen developing-country governance. Donors and international institutions can provide aid in ways that can impose practices and reporting requirements that fragment and overwhelm already fragile governance systems, or in ways that help strengthen governance. Many of the areas noted in chapter 3 and addressed in the Paris Declaration are relevant for the governance agenda. Beyond aid, global check and balance mechanisms can provide new governance instruments for helping poor countries meet the MDGs. Recent work on standards and codes provides sources of good practice for all countries, and is increasingly being used to benchmark performance. Part II of the GMR therefore considers both national governance systems and the emerging global framework to support good governance.

Notes

1. The relationship remains robust with and without the inclusion of developed countries in the sample. It remains statistically significant, though somewhat weakened, when nations in the East Asia and Pacific region also are excluded from the sample.

2. Svensson (2005), quoting Transparency International's 2004 report.

3. Svensson (2005), reporting results from Di Tella and Schargrodsky (2003); Yang (2005).

5

Monitoring Developing-Country Governance

Governance is central in development, so it is natural that many stakeholders want to monitor it. Also natural is that these stakeholders use governance measures in different ways:

- Citizens in developing countries can use measures of governance to hold governments accountable for their actions—at the micro level for the quality of service provision, at the aggregate level for the responsiveness of government action to the public interest, at all levels for the probity of using resources.
- Governments in developing countries (and development partners seeking to provide technical support) can use governance measures to improve the design of policy—for example, by providing "actionable" guideposts for operational efforts to improve governance.
- Donors in their role as funders can use governance measures for cross-country comparisons (focusing on "levels" at a point in time) or to monitor the trends within individual countries over time.

Cross-country comparisons of the level of governance are useful if the intent is to calibrate donor support according to the quality of a country's governance. Measures of trends are useful if the intent is to support "turn-

around" countries making serious efforts to transform their governance systems. Monitoring trends can signal whether these turnarounds are on track—or have stalled or gone into reverse.

This chapter addresses the challenge of monitoring developing-country governance. First, it presents a framework for monitoring. Second, it identifies indicators that are useful for monitoring the different parts of the framework. Third, it highlights some of the opportunities—and perils—that confront the governance monitoring exercise. Together, the framework, the indicators, and the overview of governance monitoring set the stage for more disaggregated analysis in chapter 6.

A Framework for Monitoring Governance

Monitoring developing-country governance has become a growth industry. A recent publication of the United Nations Development Programme (UNDP 2004), *Governance Indicators: A Users' Guide*, details 33 data sources and lists a further 33 that did not meet UNDP standards for inclusion. This focus on measurement has led to some important advances, but it has also underscored some difficulties.

Governance is more complex than it seems. While the word is often used as a euphemism for corruption, a country's governance system

BOX 5.1 Governance and corruption are not the same thing

Governance and corruption often are used synonymously. But they are quite different concepts—and conflating them can be very damaging.

Public sector governance refers to the way the state acquires and exercises the authority to provide and manage public goods and services—including both public capacities and public accountabilities. Viewed from the perspective of this report, the relevant aspects of governance are those for achieving the Millennium Development Goals (MDGs). This narrows the terrain somewhat, but perhaps less than it might seem at first sight, given the relationships between transparency, participation, and accountability on the one hand and performance in reducing poverty on the other.

Corruption is an outcome. It is a consequence of the failure of any of a number of accountability relationships that characterize a national governance system—from a failure of the citizen-politician relationship (which can lead to state capture) to a failure of bureaucratic and checks and balances institutions (which can lead to administrative corruption). Aggregate measures of corruption thus offer a useful overview of the degree to which the national governance system as a whole—rather than any part—is dysfunctional.

Perceptions of corruption can have a profound impact on a country's prospects. At home, they can break (or make) the reputation of political leaders—and affect civic perceptions of the legitimacy and trustworthiness of the state. Globally, these perceptions influence decisions on private capital flows and aid. Estimates of corruption raise awareness and attention, including through media focus on rankings. Even if these estimates have a high margin of error, with movements of a few points in one or another direction too small for any robust implication, they can still be useful. Yet, an exclusive focus on this outcome of a governance system has caused some countries to emphasize simple-minded (and largely failed) anticorruption initiatives—to the neglect of the complex challenge of strengthening national governance systems themselves.

comprises the full array of state institutions and the arrangements that shape the relations between state and society (box 5.1). Measurement can explore the broad consequences of how the governance system functions (with corruption as a major example). It can also focus more narrowly on the quality of the different institutions that make up a country's governance system. Conceptual clarity is needed to draw the distinctions among these different types of measures.

A national governance system includes many institutions and actors, including judges, legislators, tax inspectors, teachers, and accountants. Each needs the capacity to perform his or her function effectively. Effective governance also calls for the players to be accountable, often in complex ways. A school, for example, is potentially accountable to parents, to officials in departments of education (at local and central levels), to courts, and to politicians (again, both national and local).

To monitor governance—and to improve it—a framework is needed to cut through the complexity. Figure 5.1 illustrates one approach. It shows the key *actors* in a national governance system, and the key *accountability* relationships that align the incentives of the principals at each level with those of the agents delegated to act for them. As the figure suggests, transparency is an essential cross-cutting aspect of the governance system, contributing to the efficacy of both the actors and the accountability relationships.

For each actor the relevant capacities comprise the skills adequate to the task at hand, the organizational management systems capable of deploying human and other resources, transparent provision of the information needed for action, and the leadership to organize the various parts of the system and motivate its participants from the inside. Each accountability relationship rests on the following:

FIGURE 5.1   National governance systems: actors and accountabilities
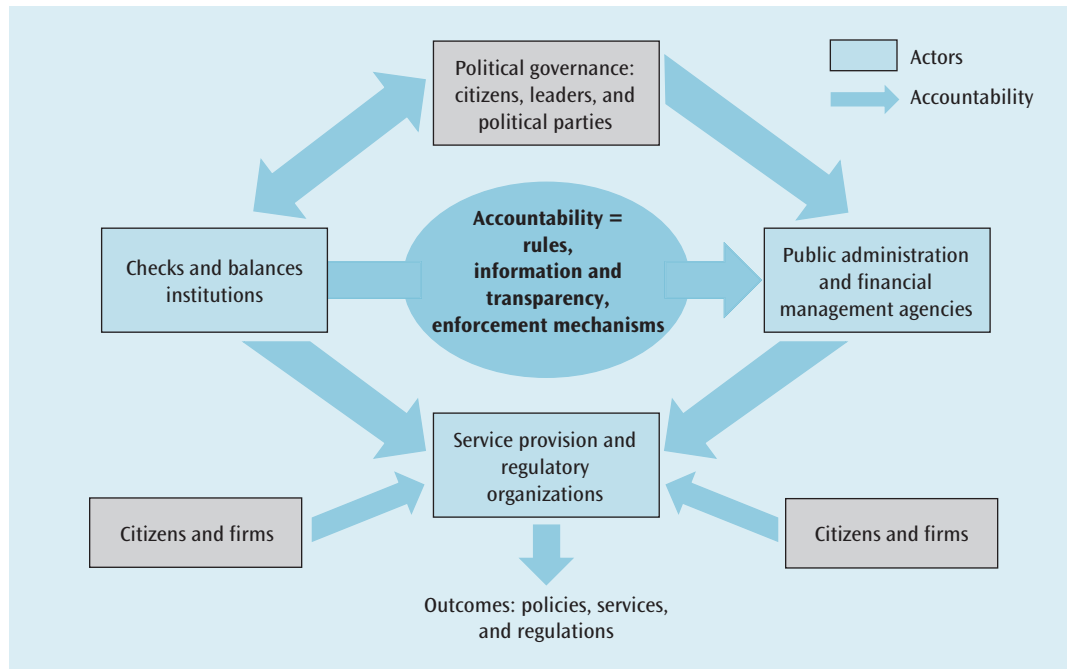
Source: Authors.

- Rules to delegate authority and indicate constraints and expected results.
- Information flows to enable principals to monitor how well agents are performing. *Transparency* of information flows increasingly is becoming the norm—on the principle that citizens always are the ultimate principals—even where the immediate principal-agent relationship might be among different parts of the state. As table 5.1 highlights, transparency's role potentially is pervasive—from the political apex of a national governance system, all the way to the service provision front line—creating plentiful opportunities for engagement.
- Enforcement mechanisms that reward successful performance and sanction those who fail to perform well.

Development outcomes depend on the operation of the national governance system as a whole. Sustained good results imply that the capacities of the players and the accountabilities among them are strong.

Turning to the five sets of actors identified in figure 5.1, the first comprises citizens, leaders, and political parties in the political governance subsystem—that is, the mechanisms for citizens to select their political leaders at national and subnational levels and the ways organized groups of citizens influence politics and government.

Politics is the prime influence on governance. Political leaders set the objectives for the rest of the governance system. Sometimes politicians work to address the general interest. Other times, their behavior is clientelistic, in the sense that "even though the average citizen is poor, politicians . . . shift public spending to cater to special interests, to core supporters, or to 'swing' voters" (World Bank (2004a: 80). What shapes whether politicians behave in developmental rather than clientelistic ways? Only rarely can they decide themselves—usually they are constrained by having to maintain their support base. Skillful, farsighted politicians, especially those at the head of political movements rooted in a

TABLE 5.1 Nodes of transparency in national governance systems

Transparency node	Specific examples of transparency
Node 1: Transparency in political governance	■ Disclosure by political candidates and public officials of assets, education, and criminal record ■ Widely available public information on the performance of government
Node 2: Transparency in public administration and financial management	■ Open competitive system of public procurement ■ Meritocratic recruitment of public sector employees ■ Transparent in-year financial reporting
Node 3: Transparent oversight of public administration by checks and balances institutions	■ Participatory budget formulation process, including cabinet-level and parliamentary discussions ■ Timely, comprehensive, high quality, and publicly available audit of budget by independent institutions
Node 4: Checks and balances institutions ensure open flow of information	■ Free press ■ Freedom of information act ■ Publication of judicial and administrative decisions ■ Open decision-making processes ■ Publication of parliamentary debates
Node 5: Transparency in relationship between citizens/firms and service providers	■ Independent service delivery scorecards/surveys ■ Public information on results by provider organizations (monitoring and evaluation data, annual reports, and so on) ■ Service charters issued by provider organizations ■ Publicly posted information on financial and other resources provided to provider organizations

Source: Authors.

broad social vision, may be able to shape the objectives for their supporters. But the process is often driven more by responding to the interests of key allies, and sometimes these powerful interests may capture the state. Note that a democratic electoral process does not guarantee that politicians will focus on the general interest, and that the general interest can also guide the action of some politicians in nondemocratic societies.

A second set of actors consists of checks and balances institutions, including parliaments, independent oversight agencies (supreme audit institutions, ombudsmen, anticorruption commissions), the judicial system, a free press, and democratically accountable local institutions. These institutions have at least three distinct functions. They establish the rules of the game for political competition. They provide the rules of the game for the broader working of

civil society and for the operation of the private market economy. And they limit the influence of politicians on the bureaucracy. Checks and balances institutions are a ubiquitous feature of polities—not only of liberal democracies—though their specific forms can vary with the mode of political organization.

The third set of actors comprises the cross-cutting control agencies responsible for public administration and financial management, including those responsible for budget formulation, execution, and reporting systems; procurement systems; monitoring and evaluation systems; intergovernmental systems; and civil service management systems. The fourth set includes service provision and regulatory organizations, including sectoral line ministries, autonomous public or private frontline providers, and regulatory agencies. These two sets of players make up the public

bureaucracy, which follows objectives set by political leaders within a framework set by checks and balances institutions. The justice system plays a dual role, as a checks and balances institution and a provider of dispute resolution services to society.

Within the bureaucracy, cross-cutting bureaucratic control agencies oversee service provision and regulatory agencies. This is where options for prioritization and resource allocation are developed—and where the responsibility lies for establishing and enforcing the rules and accountabilities (for example, the financial management and personnel rules) within which service provider and regulatory agencies operate.

The fifth set of actors comprises citizens and firms—as users of public services, including regulatory services and service providers. Citizens and firms can be depicted as principals, holding providers (agents) accountable for the efficiency and effectiveness of service provision. The extent to which they can do this depends on the quality and transparency of information flows. In markets with both public and private providers, competition can be a powerful disciplining influence on all providers—including those in the public sector.

To illustrate the relevance of the framework, one might consider two very different configurations along a spectrum of governance quality. The first pattern—sometimes termed *good enough governance*—summarizes the attributes of consistently stronger governance.[1] For governance to be good enough, the public bureaucracy need not perform at the highest levels of efficiency. More important is that the accountability arrangements built into the national governance system be mutually reinforcing, so that the system can self-correct. Failure in one part of the system (such as corruption in the use of public funds) generates pressures from other parts (parliament, courts, or citizen groups) to refocus on the public purpose.

The contrasting pattern—*clientelist governance*—characterizes countries with much weaker governance performance. In clientelist countries, formal and informal systems of authority work at cross-purposes, and the latter dominates the former.[2] Political leaders use their control over patronage resources to maintain their power base; at the limit, they are captured by powerful private interests. Leaders can bypass or override checks and balances institutions and the public administration when these get in the way of their political goals. Systems are not transparent. Levels of corruption are generally high. Informal norms are, of course, also a reality in better-governed settings; however, they do not conflict as egregiously with the formal arrangements.

Clientelist systems may limit development, but they can be stable if political leaders choose to exercise sufficient restraint to enable the formal system to operate, however imperfectly. This is more likely when they take a long-term view and recognize the importance of sustaining the institutional capacity to govern. The result—as observed in Africa in the 1970s and 1980s, for example—may be a seemingly long-term clientelist equilibrium. But this equilibrium can turn into an accelerating downward spiral if the time horizon of leaders is short. Bureaucratic decay deepens as organizations lose resources and competent staff. Economic decay deepens as public services weaken and policy becomes more capricious. Investor confidence evaporates and political decay deepens as the leadership finds itself trying to buy off constituencies with fewer and fewer resources. At the limit (as in Sierra Leone) the endpoint of this downward spiral of decay can be the collapse of the state.

For many low-income countries, improving governance means breaking out of the trap of clientelism. Because clientelism (as is the case with all governance arrangements) is deeply intertwined with the structure and exercise of political power, this can be enormously difficult. Different societies find different ways to break free. As a result, their trajectories of governance reform vary—with corresponding differences across countries as to which actors and accountabilities improve rapidly, and which lag. As will become evident, these variations have important implications for both governance monitoring and reform.

Indicators of Governance

Even with greater clarity about the relevant institutions, measuring their quality is difficult. Formal systems can be categorized and rated—but the gap between formal arrangements and realities on the ground is often very wide. Institutional processes are difficult to observe and measure systematically. Some outcomes can be measured, but these can have multiple causes and are often remote from the quality of governance. There are ways of responding to these measurement difficulties, but all are imperfect—and often subject to large margins of error.

The framework suggests two distinct approaches to monitoring the quality of governance. The first is to monitor at a disaggregated level, using specific measures of the quality of key governance subsystems—and to use the results as "actionable indicators" to identify specific strengths and weaknesses in individual countries, and thus to guide reforms and track progress. The second is to monitor governance at more aggregate levels, using broad measures. Broad measures have different uses from their specific counterparts. They can help reveal some systematic patterns underlying the complexity and diversity across individual subsystems. And they can provide some basis for monitoring overall trends across countries and over time.

Broad measures of governance can be derived in two ways. First, they can be composite measures built up from disaggregated indicators. Sometimes, such measures are constructed in a way that makes it possible to drill down from aggregate to disaggregated levels, and thereby identify strengths and weaknesses in an action-oriented way. In practice, governance measurement has not yet advanced to the point that this is routinely feasible—despite some advances in this direction. Second, broad indicators could be derived by focusing on the outcomes produced by national governance systems.

This report reviews and applies both broad and specific governance indicators. Irrespective of the level of aggregation, few governance data sets are objectively measurable indicators: multiple attempts reveal how difficult it is to construct them.[3] The data mostly reflect subjective perceptions—sometimes expert assessments, other times survey-based measures of the perception of citizens or firms. Some surveys, however, do ask questions that produce "objective" data—for example, the share of household income or sales revenues used to pay bribes. While the use of expert assessments and perception-based data is ubiquitous in the social sciences, caution—and careful attention to the likely margins of uncertainty—is needed in the interpretation of results. Indeed, as the next sections will illustrate, some measurement error is inevitable, regardless of the type of governance measure used.

The Variety

Few of the 33 data sources listed in *Governance Indicators: A Users' Guide* (UNDP 2004), can be used straightforwardly in this report. Some fall short of the requisite comprehensiveness of country coverage—particularly that of low-income countries. Others are collected irregularly, weakening their ability to measure trends. This report focuses on a subset of 14 measures that offer comprehensive country coverage and, among them, cover each of the diverse facets of national governance systems in the framework. (See table 5.4 at the end of this chapter.) Three sets of indicators—the World Bank's Country Policy and Institutional Assessments (CPIAs), which account for 5 of the 14 indicators; the Kaufmann-Kraay aggregate governance indicators, which account for 3 additional indicators; and 3 selected indicators from the Doing Business database and the Investment Climate Surveys—are examined in this chapter. Other indicators are examined in chapter 6.

The CPIA

In the late 1970s the World Bank began using systematic country assessments to

guide the allocation of International Development Association (IDA) resources. By the late 1990s the CPIA had evolved to something close to its current format. A further round of fine-tuning came in 2004, to implement suggestions by an independent panel of outside experts. The 2005 CPIAs will be the first publicly available detailed scores for IDA countries.

CPIAs examine policies and institutions, not development outcomes, which can depend on forces outside a country's control. The CPIA looks at 16 distinct areas grouped into four clusters (box 5.2). For each criterion, very detailed guidelines are provided to help Bank staff score individual countries along an absolute 1–6 scale.

For monitoring governance systems, the CPIA indicators can be used in three ways:

- At the most disaggregate level, scores for individual criteria in the public sector management and institutions cluster can be quite specific, and actionable. The analysis of public budget and administrative management systems in the next chapter will draw on CPIA criteria 13 and 15.
- The average score for the public sector management and institutions cluster (cluster D) can be an aggregate indicator of the quality of a country's governance system. Table 5.2 places 66 low-income, potential IDA-recipient countries into five groups, according to their CPIA cluster D scores for 2004.[4]
- The average score for clusters A, B, and C can be an aggregate measure of the quality of a country's economic and sectoral policies—viewed as an outcome measure of the quality of a country's governance system.

As with all indicators, the CPIA has its limitations. The assessments are made by World Bank staff. Even if expert in their field and well informed about individual countries, staff sometimes may not be aware of the intimate details as to how things really work in a country. Some of the criteria do not lend themselves readily to an ordinal scale of qual-

---

BOX 5.2   The 2004 CPIA's 16 criteria

A. **Economic management**
  1. Macroeconomic management
  2. Fiscal policy
  3. Debt policy
B. **Structural policies**
  4. Trade
  5. Financial sector
  6. Business regulatory environment
C. **Policies for social inclusion/equity**
  7. Gender equality
  8. Equity of public resource use
  9. Building human resources
  10. Social protection and labor
  11. Policies and institutions for environmental sustainability
D. **Public sector management and institutions**
  12. Property rights and rule-based governance
  13. Quality of budgetary and financial management
  14. Efficiency of revenue mobilization
  15. Quality of public administration
  16. Transparency, accountability, and corruption in the public sector

---

ity—even though "the criteria were developed to ensure that, to the extent possible, their contents are developmental neutral, that the higher scores do not set unduly demanding standards, and can be attained by a country that, given its stage of development, has a policy and institutional framework that strongly fosters growth and poverty reduction" (World Bank (2004b: 5). Policy expectations in some areas, such as social protection, are different for low- and higher-income countries. Staff assessments can be affected by the fact that the CPIA forms the basis for allocating IDA resources. There are risks of ideological bias—for example, on the merits of low tariffs versus an export-neutral combination of tariffs and subsidies. To address these limitations and ensure consistency across countries, the World Bank goes through an elaborate multistage process for scoring the CPIAs. The process includes an initial round of benchmarking by a global

TABLE 5.2   2004 country scores for the CPIA public institutions cluster

CPIA institutions cluster score	Countries
Above 3.5	Armenia, Bhutan, Ghana, India, Mali, Senegal, Tanzania
3.3–3.5	Benin, Bosnia and Herzegovina, Burkina Faso, Ethiopia, Georgia, Honduras, Indonesia, Kenya, Lesotho, Madagascar, Malawi, Nicaragua, Pakistan, Rwanda, Serbia and Montenegro, Sri Lanka, Uganda, Vietnam
3.0–3.2	Albania, Azerbaijan, Bangladesh, Bolivia, Cameroon, Eritrea, Guyana, Moldova, Mauritania, Mongolia, Mozambique, Nepal, Niger, Papua New Guinea, São Tomé and Principe, Zambia
2.6–2.9	Burundi, Chad, Rep. of Congo, Côte d'Ivoire, Djibouti, The Gambia, Guinea, Kyrgyz Republic, Nigeria, Sierra Leone, Solomon Islands, Tajikistan, Uzbekistan, Republic of Yemen
2.5 or below	Angola, Comoros, Central African Republic, Dem. Rep. of Congo, Guinea-Bissau, Lao PDR, Sudan, Togo

Source: World Bank.

team drawn from across the World Bank, subsequent rounds within operating regions using the benchmarked countries as guideposts, and a further round of validation by central units. The results are discussed with national governments, but final scoring rests with the Bank.

Given the above, it is appropriate to interpret CPIA scores as estimates, with some margin of error. Estimates place the standard error at about 0.24 for aggregate measures on the 1–6 scale (see Gelb, Ngo, and Ye 2004). High-, middle-, and low-performing countries can thus be distinguished, but attempts to position countries on a fine scale will likely result in some being misclassified. Small annual changes may not be easy to assess with confidence but CPIA-type estimates are able to distinguish trends—for 1999–2004 the correlation between the CPIA and the similar index produced by the African Development Bank is 0.8, with changes moving in the same direction for 43 out of 51 countries. Ultimately open debate offers the best way of uncovering and addressing remaining weaknesses in the CPIA. The decision to make public the detailed 2005 CPIA scores for IDA-recipient countries is an important step in the ongoing process of enhancing the transparency of this potentially important indicator.

Kaufmann-Kraay

The Kaufmann-Kraay (KK) indicators, published on the Web site of the World Bank Institute, are the product of research conducted by World Bank staff. But unlike the CPIAs, they are not a formal World Bank product, and they are not used in any systematic way in World Bank decisions. They are one response to the problem of aggregation. They generate a set of six composite aggregate indicators from a proliferation of loosely connected disaggregated measures by using a technique for statistical aggregation, the unobserved components model. The six aggregate governance indicators are as follows:

- Voice and accountability
- Political stability and absence of violence
- Government effectiveness
- Regulatory quality
- Rule of law
- Control of corruption

These indicators, available for every second year from 1996 to 2004, are an amalgam of 352 variables, culled from 37 data sources produced by 31 organizations. Three of these indicators—voice and accountability, rule of law, and control of corruption—will be used as part of governance monitoring in this and the next chapter.

Kaufmann and Kraay explain the aggregation approach as follows:[5]

The premise [is that] each of the individual data sources provides an imperfect signal of some deep underlying notion of governance that is difficult to observe directly . . . that, within each cluster, each of these indicators measures a similar underlying basic concept of governance . . . [The challenge is to] isolate the informative signal about governance from each data source, and to optimally combine the many data sources to get the best possible signal of governance in a country based on all the available data . . . The unobserved components model expresses the observed data in each cluster as a linear function of the unobserved common component of governance, plus a disturbance term capturing perception errors and/or sampling variation in each indicator.

The aggregation procedure used by KK has some important strengths for empirical work on governance. The methodology enables very broad country and territory coverage—209 in the most recent version. The aggregation procedure generates, for each country, both point estimates and standard errors of these estimates.

By making explicit the standard errors of their estimates, the KK measures highlight how challenging it is to measure precisely the quality of governance—for both broad and specific measures. Specific measures often are based on sample surveys or on expert assessments—with the risk of sampling error in the former (although robust sampling and statistical methodologies can reduce the range of error), and of informant error in the latter (although robust peer review can limit this risk). By combining multiple sources into a single measure, KK helps reduce uncertainty of this type; their approach takes advantage of the well-known statistical property that the margin of error of a measure declines as the number of independently generated estimates increases.

The cost, though, is to introduce a different type of uncertainty. KK's gain in *precision* is offset by a loss of *specificity*. The KK composite variables combine sources that measure similar, but not identical, phenomena. A country's poor score on one of the aggregate indicators might reflect divergent performance among the underlying sources, making the indicators less useful as actionable tools. Further, the KK methodology both presumes that each of its sources is independently generated, and weights converging sources more heavily. But if in fact the converging sources draw on a shared underlying model, the methodology risks marginalizing sources that offer a view that is different—but not necessarily wrong.[6]

The example of corruption illustrates the use and limitations of aggregate indicators. Two sets of aggregate indicators—the KK "control of corruption" aggregate indicator and Transparency International's (TI) Corruptions Perceptions Index[7]—provide "best practice" broad measures, and so have dominated cross-country ratings of corruption. Both rank countries according to their perceived performance in controlling corruption, and both report margins of error of their estimates; the country estimates for both are included in the statistical annex. Each indicator draws on multiple primary indicators to produce country rankings. The KK indicator description indicates what is being measured, namely:

. . . perceptions of corruption, conventionally defined as the exercise of public power for private gain. The particular aspect of corruption ranges from the frequency of "additional payments to get things done" to the effects of corruption on the business environment, to measuring "grand corruption" in the political arena or in the tendency of elites to engage in "state capture . . .

Although the KK and TI methodologies differ,[8] in practice their results are very similar, with a correlation coefficient of 0.97. Either indicator can be used for comparisons across countries and over time. Two conclusions:

■ High-income Organisation for Economic Co-operation and Development (OECD) countries are usually in the least corrupt third of countries. The distribution of middle-income countries is mixed; at least 8, and (depending on the confidence interval used) up to 17 middle-income countries are in the bottom third. The majority of low-income, IDA-recipient countries are in the most corrupt third, though 20 countries in this group are in the middle third. This distribution reflects the well-known inverse correlation between per capita income and levels of corruption. But even within the general pattern, levels of corruption vary widely for countries at similar levels of per capita income—implying that corruption is not wholly tied to income.

■ It is possible to position countries on a global corruption spectrum only in broad terms. Even if countries are organized into three broad categories—top, middle, and bottom thirds—only for about 40 percent of countries, and only for half of those in the bottom third, can one have at least 95 percent confidence that they belong in the third in which they have been placed.

Doing Business and the Investment Climate Surveys

The World Bank-sponsored *Doing Business* (DB) and *Investment Climate Surveys and Assessments* (ICS) were introduced in chapter 1. These surveys are designed to monitor the business environment, not governance. Some of the business environment measures can nonetheless be directly linked to governance, and therefore are useful for governance monitoring.

The DB and ICS methodologies are very different. The ICS captures business perceptions on the biggest obstacles to enterprise growth, the relative importance of various constraints to increasing employment and productivity, and the effects of a country's investment climate on its international competitiveness. DB indicators comprise detailed, objective measures of the time and cost of strict compliance with government regula-

tions affecting private business across 10 topic areas—the number of procedures required to accomplish the task in question, the number of days necessary to accomplish the task, and the monetary cost in required fees. Input and verification is provided by government officials, lawyers, business consultants, accountants, and other professionals administering or advising on regulatory requirements.

Both the DB and ICS data are, of course, subject to measurement error. Also, differences in their methodologies generate different strengths and weaknesses, making them usefully complementary to one another:

■ *De jure versus de facto*: The DB product measures the de jure business environment whereas the ICS product measures the de facto business environment. Both measures are useful, though it is important not to confuse changes in the de jure environment with actual changes on the ground.[9]

■ *In-depth versus holistic perspective*: The DB product zeroes in on a narrow set of transactions, which it presumes to be illustrative of the business environment more broadly, whereas the ICS product provides a holistic view of the business environment from the perspective of firms themselves.

■ *Cost and coverage*: DB covers 155 countries, and all country scores are updated annually. The ICS is a more effort-intensive product than the DB product. Its database contains information on about 60 countries; it aims to cover 20–30 countries each year and resurvey each country every three years or so. High costs somewhat limit the ICS' usefulness as a tool for ongoing governance measurement across a large number of countries.

A first potential use of these business environment indicators for governance monitoring is as overall outcome measures, complementing the CPIA and KK. Box 5.3 highlights three specific measures which can play this role: the ICS measures of corruption, plus two measures of transactions costs associated with bureaucratic red-tape.

> ## BOX 5.3 Three aggregate governance Doing Business and Investment Climate Survey indicators
>
> **Unofficial payments for firms to get things done (percentage of sales) (ICS)**
>
> Average value of gifts or informal payments to public officials to "get things done" with regard to customs, taxes, licenses, regulations, services, and the like. The values shown indicate a percentage of annual sales.
>
> **Dealing with licenses (DB)**
>
> The number of procedures, average time spent during each procedure, and official cost of each procedure involved in obtaining necessary licenses and permits, completing required notifications and inspections, and obtaining utility connections (using construction of a warehouse as a benchmark example).
>
> **Senior management time spent dealing with requirements of regulations (percent) (ICS)**
>
> Average percentage of senior management's time that is spent in a typical week dealing with requirements imposed by government regulations (such as taxes, customs, labor regulations, licensing, and registration), including dealings with officials, completing forms, and the like.

Irrespective of how much business regulation a country judges to be appropriate, it always is developmentally desirable to minimize the time and hassle spent in complying. In general, high transactions costs signal some combination of an unresponsive bureaucracy, or a clientelistic environment geared to provide opportunities for informal rent-extraction by public officials.

A second potential use of the ICS data in particular is to distinguish among different types of corruption. Corruption sometimes is disaggregated into two basic forms—*state capture* and *administrative corruption*. State capture refers to the actions of individuals, groups, or firms in either the public or private sectors to influence the *formulation* of laws, regulations, decrees, and other government policies to their own advantage as a result of the illicit and nontransparent provision of benefits to public officials. State capture is commonly found in states that control important national assets, either through ownership (for instance, mineral rights, state-owned enterprises) or regulation (for instance, economic or environmental), but have limited political competition

and weaker checks and balances. Administrative corruption refers to the provision of illicit and nontransparent benefits to influence how these established rules are *implemented*. Administrative corruption flourishes in states with weaker bureaucratic capacity and accountability. The ICS surveys conducted in Eastern and Southeastern Europe and the former Soviet Union (known as the BEEPS surveys)[10] framed their questions in a way that made it possible to distinguish among the different types of corruption. The surveys found that the relative balance between state capture and administrative corruption can vary widely from country to country.

A third use of the DB and ICS data is to measure trends over time for specific features of the governance environment within individual countries. Where survey variables are narrowly defined, and the sampling and statistical work are careful, the resulting measures can have small margins of error, and so detect incremental changes over time. Figure 5.2 draws on the 1999, 2002, and 2005 BEEPS surveys to report trends in administrative corruption for the 26 surveyed countries.

FIGURE 5.2   Administrative corruption in Europe and Central Asia

a. Bribe frequency

b. Bribe tax

c. Corruption as an obstacle to doing business
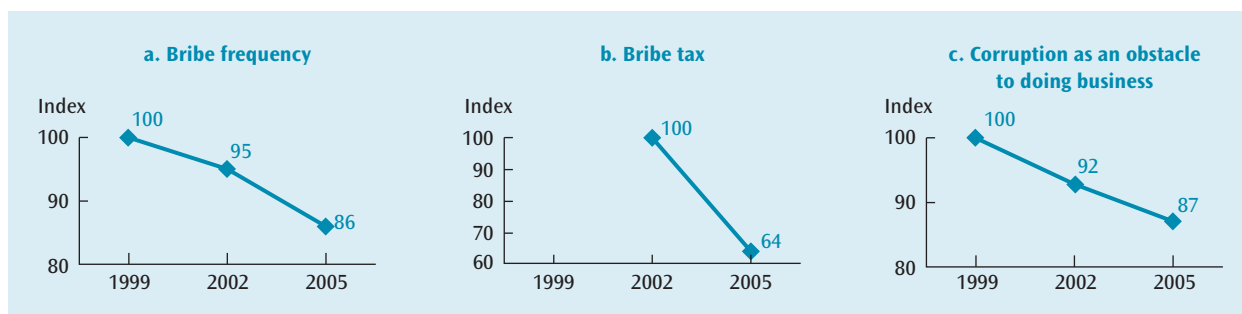
Source: The World Bank and the EBRD Business Environment and Enterprise Surveys (BEEPS) 1999, 2002, 2005.
Notes: The charts depict 2002 and 2005 values relative to 1999, except in the case of bribe tax, where the 2005 value is shown relative to 2002. Due to a change in wording, the bribe tax is not comparable between 1999 and 2002. Values are based on the simple average of country means over all countries that were present in all years.

FIGURE 5.3   Corruption in specific sectors in Europe and Central Asia, 2002–5
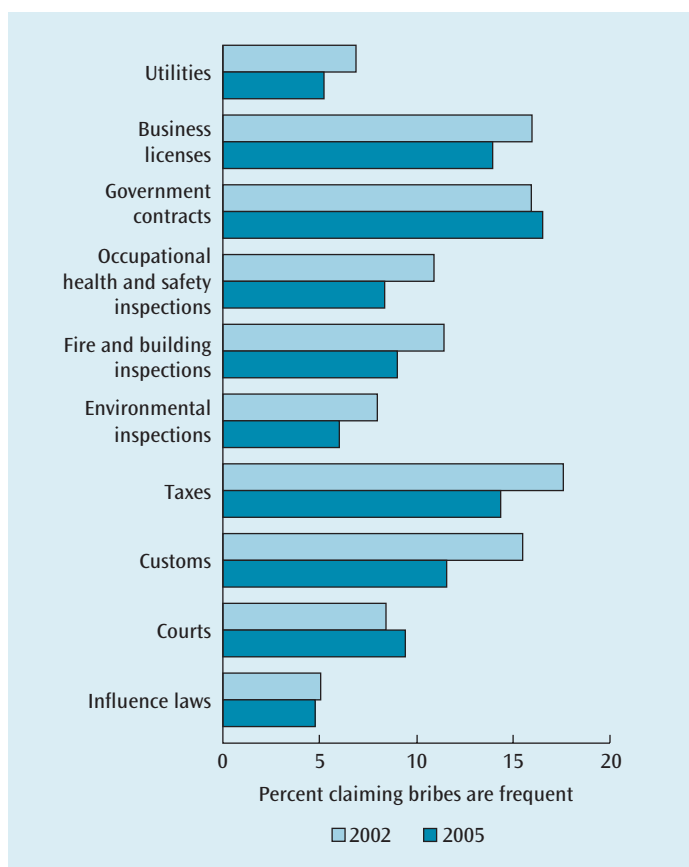
Source: The World Bank and the EBRD Business Environment and Enterprise Surveys (BEEPS) 2002 and 2005.
Notes: The chart depicts the simple mean of the country averages of the percent of firms that said bribes were frequent.

The figure points to broad declines in administrative corruption—an important success, and one that needs careful and sustained survey analysis to become evident.[11]

Finally, the DB and ICS data also are useful as specific, disaggregated measures of the performance of individual public service provision or regulatory agencies or sectors. The annex table in chapter 6 highlights four categories of specific measures that are useful for governance monitoring: measures of corruption; measures of transactions costs associated with red tape; measures of the quality of provision of education, water, and telecommunications services; and measures of the quality of justice and the rule of law. Figure 5.3 illustrates the products' usefulness with data (again for the 26 BEEPS survey countries) on variations across sectors in the incidence of corruption.

Governance Monitoring—From Broad to Specific

Ten years ago, none of the Transparency International, KK, and Doing Business indexes existed; the CPIA was still quite rudimentary, and none of it was public; and ICS were not being done on a systematic basis globally. Over the past decade, there have thus been major advances in the development of broad indicators for monitoring developing-country governance.

Combining different indexes yields some striking patterns, additional to those described above. Table 5.3 combines the country results for two sets of broad outcome indicators—the KK control of corruption measure, and the quality of a country's economic and sectoral policies (as measured by the average score for CPIA clusters A, B, and C).[12] Three key lessons emerge.

First, the broad indicators offer a first approximation of the patterns of variation in governance performance among 66 IDA recipient countries. About one-third of countries generally are in the higher "good enough" quintiles of both broad governance outcome indicators. Another third are lodged firmly in the lower "clientelistic" quintiles of the indicator sets.[13]

Second, these consistent clusters aside, what is especially striking is the uneven mix of strengths and weaknesses for individual countries. Country performance is broadly similar on both dimensions for 34 of the 66

IDA-eligible countries. But (even making substantial allowance for the margins of error in measurement) in 17 countries the quality of policy and institutions is better than performance on corruption, and 15 countries show the opposite pattern. Bangladesh currently is perhaps the best known example of a country with relatively weak perceived control of corruption but strong performance on policies and on poverty reduction—though, as table 5.3 suggests, many other countries evince a similar pattern.

Countries can thus differ. Some with weak policies appear to be less corrupt. And others, stronger on the policy front, seem less successful in controlling corruption. This divergence raises some questions, both for governance monitoring and more broadly:

■ If the divergence is not simply the result of measurement error (and the allowance for large margins of error in the construction of table 5.3 suggests that it mostly is not),

TABLE 5.3   Intermediate outcomes—corruption versus policy

Relative performance across governance outcomes	CPIA 2004 policy quintiles (cluster (a)–(c) average)				
	Bottom quintile	4th quintile	3rd quintile	2nd quintile	Top quintile
Control of corruption and policy performance are broadly similar[a]	Angola, Central African Republic, Comoros, Dem. Rep. of Congo, Côte d'Ivoire, Lao PDR, Nigeria, Solomon Islands, Sudan	Burundi, Cambodia, Rep. of Congo, Djibouti, Papua New Guinea, Sierra Leone, Zambia	Cameroon, Ethiopia, Kenya, Malawi, Moldova, Nepal, Rwanda, Mozambique, Niger, Yemen	Benin, Bosnia and Herzegovina, Mali, Serbia and Montenegro, Sri Lanka	Burkina Faso, Nicaragua, Senegal
Better policies, weaker control of corruption[b]		Chad, Haiti, Uzbekistan	Tajikistan	Bangladesh, Georgia, Kyrgyz Republic, Indonesia, Vietnam	Albania, Armenia, Azerbaijan, Bolivia, Honduras, Pakistan, Tanzania, Uganda
Better corruption control, less effective policies[c]	Eritrea, Guinea-Bissau, São Tomé and Principe, Togo, Zimbabwe	The Gambia, Guinea, Mauritania	Mongolia, Lesotho	Bhutan, Ghana, Guyana, India, Madagascar	

Source: World Bank CPIA Database.
a. Country percentile rank for the CPIA Policy Outcome and the Kaufmann, Kraay, and Zoido-Lobaton (KKZ) Control of Corruption indicator are less than 20 percentile points apart.
b. Country percentile rank for Policy Outcome is better than Control of Corruption by percentile rank of at least 20 points.
c. Country percentile rank for Control of Corruption is better than Policy Outcome by percentile rank of at least 20 points.

is it pointing to some underlying structural or social features, or to differences across countries in the relative importance ascribed to good policy on the one hand, and the fight against corruption on the other?

- How do differences in the relative importance ascribed to these different dimensions affect country performance on poverty reduction?
- How should donors respond to this divergence? Should they differentiate their support as between countries that fight corruption determinedly but have relatively weak policy, and countries stronger on the policy front than on reining in corruption?
- How do such differences reflect the performance of specific governance subsystems? Which are most relevant for containing corruption?

The third lesson on indicators (broad and specific) follows from the fact that they generally have large errors. Quantifying these margins of error has been an important advance over the past decade. The consensus among researchers is that, by and large, the broad governance indicators we have are what we will have to work with—no breakthrough capable of providing an overarching, yet precise measure of governance is on the horizon. This signals the limitations of efforts to classify countries according to their broad governance performance. Further, country-specific operational work also needs indicators that are specific, and identify "actionable" entry points for reform.

Governance monitoring thus needs to make balanced use of both broad and more specific indicators. The Global Monitoring Report (GMR) identifies 14 indicators—both broad and specific, all of which are italicized in table 5.4—as core for governance monitoring. (The CPIA, KK, DB, and ICS indicators have already been introduced; the others will be introduced in chapter 6.) Most come from sources that are updated every year or two, such as the CPIA, KK, and TI. Country coverage already is comprehensive for these and for Doing Business and Polity IV; many countries are also being included in periodic ICS. The exception is the PEFA indicator set, developed by the Public Expenditure and Financial Accountability global program. Effective monitoring of the

TABLE 5.4 Governance monitoring indicators

	Indicators with comprehensive country coverage	Other key indicators
Overall governance performance	*1, 2, 3. Control of corruption* (KK, TI, ICS) *4. Policy outcome* (CPIA cluster a–c average) *5. Aggregate public institutions* (CPIA cluster d) *6, 7. Business transactions costs* (DB, ICS)	
Bureaucratic capability	*8. Budget/financial management* (CPIA-budget) *9. Public administration* (CPIA-admin) Doing Business indicators Investment Climate Surveys Statistical Capacity	*14. PEFA indicators* Procurement "Actionable" public administration Service-provision-specific
Checks and balances institutions	*10. Voice and accountability* (KK) *11, 12. Justice and rule of law* (KK, CPIA-rules) *13. Executive constraints* (Polity IV)	Global Integrity Index

Source: Authors.
Note: Each indicator set is described in the text; the 14 italicized indicators are considered key by the GMR.

quality of public financial management is central to the new approach for scaling-up aid. *The GMR thus recommends that priority attention be given to more systematically applying the PEFA indicators in aid recipient countries.*

Along with the 14 core indicators, table 5.4 also identifies a variety of other (nonitalicized) indicators that are useful for global monitoring. All of these are specific "actionable" indicators. Of the specific indicators, the DB, ICS, and Statistical Capacity indexes already are available comprehensively. Using the others for monitoring remains a work in progress. The Global Integrity Index has been measured in 25 countries. The procurement index—the OECD-DAC [Development Assistance Committee] Baseline Indicator Set (BIS) for Procurement—has been proposed to be piloted in 10 countries. Specific "actionable" indicators measuring key aspects of public administration have been piloted in three countries. The number of service-provision-specific indicators potentially is large, though so far only one that is clearly governance related—teacher absenteeism—has been collected systematically, and so far for fewer than a dozen countries. Though the cost of developing and applying these indicators across a large number of countries is high, the PEFA and DB experiences suggest that the benefits can be higher still. *The GMR thus recommends that support be given for the further development of actionable indicators.* Once the PEFA indicators have been rolled out systematically, priority attention could be given to the expansion of coverage of the BIS procurement measures and the Global Integrity Index.

Work on these specific indicators is emerging as the frontier challenge for governance monitoring. They focus on a narrow target for measurement, so—if the indicators are carefully defined, and the methodologies for measurement robust—specific measures can provide quite tight margins of error, even if they cannot easily be used as proxies for broader governance outcomes. The narrow focus of specific measures also makes them actionable—in the sense that they can help identify specific governance weaknesses and monitor progress of reform efforts. The next chapter will focus on the monitoring of specific governance subsystems, highlighting the potential uses, both for monitoring and for governance reform, of actionable indicators.

Notes

1. For a detailed development of the term, *good enough governance,* see Grindle (2004).

2. For some theoretical and applied analyses of clientelism, see Bratton and Van Der Walle (1998); Carothers (2002); Levy and Kpundeh (2004); Lewis (1996); Migdal (1988); North (1990); Olson (1991); World Bank (2004a).

3. The Public Expenditure and Financial Accountability (PEFA) indicators incorporate a few objective measures of the quality of budget performance.

4. There are a total of 81 countries eligible for IDA in FY06 (excluding Iraq and Kosovo). Ten "small island economy exception" countries are excluded from the sample. Afghanistan and Timor-Leste do not have CPIA scores and are excluded from this sample. Liberia, Somalia, and Myanmar are inactive IDA countries, and do not have recent CPIA scores.

5. Kaufmann, Kraay, and Mastruzzi (2005: 7); Kaufmann, Kraay, and Zoido-Lobaton (1999: 9).

6. Note that, on average, composite variables do not change much with different weighting schemes.

7. For details about this index and annual results for 1995–2005, see www.transparency.org/ policy_and_research/surveys_indices/cpi.

8. See, for example, Kaufmann, Kraay, and Mastruzzi (2005), which contrasts the methodologies used by the two indexes for calculating standard errors.

9. For an interesting analysis of the relation between de jure and de facto measures of the business environment see Kaufmann, Kraay, and Mastruzzi (2005). The authors find the correlation between de jure and de facto measures to be about 0.4. The gap is larger in countries with higher perceptions of corruption, signaling the power of informality in working around de jure constraints.

10. These surveys were conducted jointly by the World Bank and the European Bank for

Reconstruction and Development; they are known as the Business Environment and Enterprise Performance Surveys (BEEPS).

11. This forthcoming analysis is tentatively titled "Anticorruption in Transition 3." The second installment of the monitoring exercise, focusing on the period between 1999 and 2002, was presented in Gray, Hellman, and Ryterman (2004).

12. An alternative approach might have been to view corruption not as an outcome but as a proxy for the overall quality of public institutions. The correlation coefficient between CPIA cluster D and the KK control of corruption measure is 0.76, suggesting that it is indeed quite a good proxy. Consistent with the substantial dispersion evident in table 5.3, the correlation coefficient between KK corruption and the CPIA clusters A–C average is 0.53.

13. The location of countries in higher and lower quintiles is broadly similar to the patterns for CPIA cluster D, in table 5.2.