Governance Matters III: Governance Indicators for 1996, 1998, 2000, and 2002

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi

Six dimensions of governance are estimated covering 199 countries and territories for four periods: 1996, 1998, 2000, and 2002. The indicators are based on several hundred individual variables measuring perceptions of governance drawn from 25 data sources constructed by 18 organizations. These individual measures are assigned to categories capturing key dimensions of governance. An unobserved-components model is used to construct six aggregate governance indicators in each of the four periods. Point estimates of the dimensions of governance are provided as well as the margins of errors for each country for the four periods. Methodological issues are also addressed, including tests for potential biases, and the interpretation and use of the data, given the estimated margins of errors for the indicators. The data and a Web-based graphical interface are available online at www.worldbank.org/wbi/governance/govdata2002/index.html.

This article presents estimates of six dimensions of governance for 199 countries and territories for 1996, 1998, 2000, and 2002 developed in the context of an ongoing project to measure governance across countries. Section I describes the data used in developing this round of the governance indicators, which include several new sources. Data sources used in the earlier studies were updated forward to 2002 and backward to 1996, and previously estimated indicators for 1998 and 2000 were revised to reflect the new data. The aggregation procedure, described in section II, provides not only estimates of governance for each country but also measures of the precision or reliability of these estimates. Although the new data have improved the precision of the governance indicators, the margins of error remain large relative to the units in which governance is measured, so that comparisons across countries and especially over time should be made with caution. Measurement error is not unique to these indicators but is pervasive among all measures of governance and

Daniel Kaufmann is director at the World Bank Institute; his e-mail address is dkaufmann@worldbank.org. Aart Kraay is senior economist in the Development Research Group at the World Bank; his e-mail address is akraay@worldbank.org. Massimo Mastruzzi is research analyst at the World Bank Institute; his e-mail address is mmastruzzi@worldbank.org. The authors thank Merli Baroudi, David Cieslikowsky, Peter Cornelius, Rui Coutinho, Guy Dunn, Rich Fullenbaum, Joel Hellman, Phil Keefer, Marta Lagos, Marc Levy, Satish Mannan, Fiona Paua, Adrian Shute, Craig Webster, and Jennifer Windsor for providing data and answering numerous questions; Steve Knack, Clay Lowery, and Steve Radelet for their comments; and Erin Hoffmann for assistance. The support and collaboration of the World Economic Forum, the U.S. State Department, the Netherlands government, and Italian Trust Funds is appreciated.

THE WORLD BANK ECONOMIC REVIEW, VOL. 18, NO. 2,

doi:10.1093/wber/lhh041 18:253–287

institutional quality. An advantage of the measures used here is that explicit margins of error reflecting this measurement error can be computed.

Section III examines issues related to the construction and use of the governance indicators, such as the usefulness of subjective measures of governance relative to alternatives. It also empirically investigates the importance of ideological biases in expert assessments of corruption, finding little evidence that they exist. To illustrate the consequences of the substantial margins of error associated with the governance indicators, the aid allocation rules proposed for the U.S. government's Millennium Challenge Account, which rely on these measures, are examined in section IV. Section V explores the limited evidence available on global trends in governance, and section VI compares the Control of Corruption indicator estimated here with the widely used Corruption Perceptions Index produced by Transparency International.

I. Measuring Governance

In this study *governance* is defined broadly as the traditions and institutions by which authority in a country is exercised. This includes the process by which governments are selected and replaced, the capacity of the government to formulate and implement sound policies, and the respect of citizens and the state for the institutions that govern economic and social interactions among them. This definition guides the construction of the governance indicators for this study.

Governance Clusters

Data on perceptions of governance from a large number of sources are organized into six clusters corresponding to six dimensions of governance. The first two clusters are intended to capture the process by which those in authority are selected and replaced. One cluster, referred to as "voice and accountability," includes indicators of the political process, civil liberties, and political rights. These indicators measure the extent to which citizens are able to participate in the selection of governments. This category also includes indicators measuring the independence of the media, which serves the important monitoring role of holding those in authority accountable for their actions. The second cluster, "political stability and absence of violence," combines several indicators measuring perceptions of the likelihood that the government in power will be destabilized or overthrown by unconstitutional or violent means. This cluster captures the idea that the quality of governance is compromised by the likelihood of a wrenching change in government that directly affects the continuity of policies and undermines the ability of citizens to peacefully select and replace those in power.[1]

1. There is some ambiguity about the normative direction of a few of the subcomponents this indicator. For example, a few sources rank countries such as Cuba and the Democratic Republic of Korea highly in terms of their political stability, which simply reflects the longevity of the governments in power.

The next two clusters summarize indicators of the ability of the government to formulate and implement sound policies. The "government effectiveness" cluster combines the quality of public service provision, the quality of the bureaucracy, the competence and independence of the civil service, and the credibility of the government's commitment to policies. The focus is on inputs the government needs to produce and implement good policies and deliver public goods. The second cluster, "regulatory quality," focuses on the policies themselves. It includes measures of the incidence of market-unfriendly policies, such as price controls or inadequate bank supervision, and perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development.

The last two clusters summarize the respect of citizens and the state for the institutions that govern their interactions. "Rule of law" includes indicators that measure how well agents abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts. Together, these indicators measure a society's success in developing an environment in which fair and predictable rules form the basis for economic and social interactions and property rights are protected. The final cluster, "control of corruption," measures perceptions of corruption, conventionally defined as the exercise of public power for private gain. The focus of the various sources differs somewhat, ranging from the frequency of additional payments needed to get things done to the effects of corruption on the business environment and "grand corruption" in the political arena. Corruption is often a manifestation of a lack of respect of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules that govern their interactions.

Sources of Governance Data

Some 250 individual measures from 25 sources produced by 18 different organizations are used in constructing the 2002 indicators (table 1; further details on each source and on how questions from each source were assigned to the six governance clusters are available online at www.worldbank.org/wbi/governance/govdata2002/index.html). These organizations include international organizations, political and business risk-rating agencies, think tanks, and nongovernmental organizations. Six new data sources are included in 2002: Afrobarometer, a survey of individuals in 12 African countries; Reporters without Borders, an assessment of press freedoms compiled by an international journalist organization; Human Rights, a numerical coding of assessments of certain dimensions of human rights as reported by the U.S. State Department and Amnesty International (first reported in Cingranelli and Richards 2001 and subsequently updated and expanded by Craig Webster); World Markets Online, a commercial risk-rating agency; Voice of the People, a citizen survey sponsored by Gallup International; and the World Bank's Country Policy and Institutional Assessment (CPIA),

TABLE 1. Sources of Governance Data

Source	Publication	Code	Type	Country coverage[a]	Representative	Availability			
						1996	1998	2000	2002
Afrobarometer	Afrobarometer Survey	AFR	Survey	12					x
Business Environment Risk Intelligence	Business Risk Service	BRI	Poll	50		x	x	x	x
Business Environment Risk Intelligence	Qualitative Risk Measure	QLM	Poll	115	x	x	x	x	x
Columbia University	State Capacity Project	CDU	Poll	98	x			x	x
Economist Intelligence Unit	Country Risk Service	EIU	Poll	115	x	x	x	x	x
European Bank for Reconstruction and Redevelopment	Transition Report	EBR	Poll	26		x	x	x	x
Freedom House	Nations in Transition	FHT	Poll	27		x	x	x	x
Freedom House	Freedom in the World	FRH	Poll	192	x	x	x	x	x
Gallup International	Gallup Millennium Survey	GMS	S	60				x	
Gallup International	50th Anniversary Survey	GALLUP	Survey	44			x		
Gallup International	Voice of the People Survey	GAL	Survey	46					x
Heritage Foundation/ Wallstreet Journal	Economic Freedom Index	HER	Poll	161	x	x	x	x	x
Institute for Management and Development	World Competitiveness Yearbook	WCY	Survey	49		x	x	x	x
Latinobarometro	Latinobarometro Surveys	LBO	Survey	17		x		x	x
Political Risk Services	International Country Risk Guide	PRS	Poll	140	x	x	x	x	x
PriceWaterhouseCoopers	Opacity Index	PWC	Survey	35				x	
Reporters Without Borders	Reporters Without Borders	RSF	Poll	138	x				x
Global Insight's DRI McGraw-Hill	Country Risk Review	DRI	Poll	111	x	x	x	x	x
State Department/ Amensty International	Human Rights Report	HUM	Poll	159	x	x	x	x	x
World Bank	Business Environment and Enterprise Performance Survey	BPS	Survey	18				x	x
World Bank	World Business Environment Survey	WBS	Survey	81	x		x	x	
World Bank	Country Policy and Institutional Assessments	CPIA	Poll	136		x	x	x	x
World Economic Forum	Global Competitiveness Report	GCS	Survey	75		x	x	x	x
World Economic Forum	Africa Competitiveness Report	GCSA	Survey	23			x		
World Markets Research Center	World Markets Online	WMO	Poll	186	x				x

[a]Countries included in the most recently available version of source.

an assessment of country performance constructed by World Bank country economists.[2]

Two of the new sources are also available for earlier years (Human Rights and CPIA). To improve the comparability of the governance indicators over time, indicators for 1998 and 2000 were revised to incorporate these sources. Two minor sources used in the past and no longer available were discarded.[3] For convenience the revised indicators are referred to as indicators for 1998 and 2000, even though the measures are based on data for a two-year period (1997–98 and 2000–01). A subset of indicators is also available for 1996. These were used to construct new aggregate governance indicators for 1996.

Two categories of sources are used: polls of experts and surveys of business-people or citizens. The choice between polls and surveys involves tradeoffs between cross-country comparability and firsthand knowledge of local conditions.[4] Polls of experts are designed to provide comparable results across countries through elaborate benchmarking. However, their reliability depends on the ability of a small group of experts to provide accurate assessments of the governance dimensions being measured.[5] Surveys typically draw on the responses of large numbers of respondents with direct knowledge of local conditions. However, to the extent that ostensibly identical survey questions are interpreted differently by respondents with different cultural or socioeconomic backgrounds, it can be difficult to make cross-country comparisons.[6]

How representative the sources are of the world as a whole is also important. A number of sources cover a large sample of developed and developing economies, whereas others cover very narrowly focused samples. Many of the poorest and smallest countries tend not to be covered by commercially oriented

2. Transparency International's Corruption Perceptions Index (CPI) is not used as a component of our aggregate corruption indicator because the CPI is itself an aggregate of sources that are already included in the corruption indicator constructed here.

3. These are the *Central European Economic Review*, which rated a sample of transition economies and ceased publication after our only use of this source in the 1997/98 indicators, and the Political and Economic Risk Consultancy, which has also discontinued its rating of a small number of Asian economies. Dropping these sources does not affect country coverage, and it makes the aggregate indicators more comparable over time.

4. For a more detailed discussion of the advantages and disadvantages of polls of experts relative to surveys of market participants, see Kaufmann and others 1999b, 2002, and Kaufmann and Kraay 2002.

5. Most of the polls of experts cover large groups of raters. For example, the Economist Intelligence Unit, based in London, draws on the views of a worldwide network of correspondents for its assessments, as does Freedom House, which is based in New York, and Reporters without Borders, based in France. Other polls of experts have a narrower institutional affiliation for their respondents. For example the European Bank for Reconstruction and Development Transition Report ratings are based primarily on the assessments of its staff in London, and the State Department component of our Human Rights measure reflects the views of U.S. State Department employees.

6. The three main sources of firm-level survey data (the Geneva-based World Economic Forum's Global Competitiveness Report, the Lausanne-based Institute for Management Development's *World Competitiveness Yearbook*, and the Washington-based World Bank's Business Environment Surveys) interview primarily domestic rather than foreign-owned firms in the countries they cover.

polls because they are relatively unattractive to foreign investors. Because there is a strong positive association across countries between governance and per capita incomes, this difference between sources makes it difficult a priori to compare indicators from sources that cover sets of countries with very different income levels. Similarly, there may be regional differences in governance that hamper simple comparisons across sources. For example, it is difficult to compare a governance rating based only on transition economies with one based on a broad set of countries. As discussed in Kaufmann and others (1999a), the methodology used here to construct aggregate governance indicators takes these differences in country coverage into account, transforming the data from individual sources into common units for aggregation across sources. Table 1 identifies whether sources are considered representative or nonrepresentative.

Aggregation Methodology

Implicit in how the data have been organized is the view that within each cluster each indicator measures a similar underlying basic concept of governance. There are considerable benefits from combining related indicators into an aggregate governance indicator for each cluster. First, the aggregate indicators span a much larger set of countries than any individual source, permitting comparisons of governance across a broader set of countries. Second, aggregate indicators can provide more precise measures of governance. Third, it is possible to construct quantitative measures of the precision (and thus margins of error) of both the aggregate governance estimates for each country and their components.

An extension of the standard unobserved-components model is used to combine the component indicators of each governance cluster into an aggregate governance indicator. The 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 or sampling variation in each indicator.[7] Thus the observed score of country j on indicator k, $y(j,k)$, is assumed to be a linear function of unobserved governance, $g(j)$, and a disturbance term, $\varepsilon(j,k)$:

$$y(j,k) = \alpha(k) + \beta(k) \cdot [g(j) + \varepsilon(j,k)], \tag{1}$$

where $\alpha(k)$ and $\beta(k)$ are unknown parameters that map unobserved governance $g(j)$ into the observed data $y(j,k)$. As a choice of units, $g(j)$ is assumed to be a random variable with mean zero and variance one. The error term is assumed to follow a normal distribution with zero mean and the same variance across countries but a different variance across indicators: $E[\varepsilon(j,k)^2] = \sigma^2{}_\varepsilon(k)$. Finally, the errors are assumed to be independent across sources: $E[\varepsilon(j,k) \cdot \varepsilon(j,l)] = 0$ for l different

7. The same methodology was used to construct previous versions of indicators; for detail, see Kaufmann and others (1999a). Unobserved components models were pioneered in economics by Goldberger (1972), and the closely related hierarchical and empirical Bayes models in statistics by Efron and Morris (1971, 1972).

from k. This imposes the identifying assumption that the only reason two sources might be correlated is that both are measuring the same underlying unobserved governance dimension.[8]

The disturbance term $\varepsilon(j,k)$ captures two sources of uncertainty in the relationship between true governance and the observed indicators. First, the particular aspect of governance covered by indicator k is imperfectly measured in each country, reflecting either perception errors on the part of experts (in the case of polls), or sampling variation (in the case of surveys). Second, the relationship between the concept measured by indicator k and the corresponding broader aspect of governance may be imperfect. For example, even if the aspect of corruption covered by indicator k (such as the prevalence of "improper practices") is perfectly measured, it may be a noisy indicator of corruption if there are differences across countries in what are considered to be "improper practices." Both sources of uncertainty are reflected in the indicator-specific variance of the error term, $\sigma^2{}_\varepsilon(k)$.

The estimate of governance for a country produced by the unobserved-components model is the mean of the distribution of unobserved governance conditional on the $K(j)$ observed data points for that country. This conditional mean is the following weighted average of appropriately rescaled scores of each component indicator:

$$(2) \qquad E[g(j) \mid y(j,1), \ldots, y(j,K(j))] = \Sigma_{k=1}^{K(j)} w(k) \cdot [(y(j,k) - \alpha(k))/\beta(k)],$$

where the weights applied to each source k,

$$w(k) = \sigma_\varepsilon(k)^{-2}/[1 + \Sigma_{k=1}^{K(j)} \sigma_\varepsilon(k)^{-2}]$$

are inversely proportional to the variance of the error term of that source. Precision weighting results in efficiency gains relative to the alternative of simply averaging rescaled scores from each source for each country. The standard deviation of this conditional distribution is also reported as an indicator of the confidence in this estimate:

$$(3) \qquad \mathrm{SD}[g(j) \mid y(j,1), \ldots, y(j,K(j))] = (1 + \Sigma_{k=1}^{K(j)} \sigma_\varepsilon(k)^{-2})^{-1/2}.$$

8. For some pairs of sources, this assumption may not be literally true. For example, it will be violated if different risk-rating agencies base their assessments on the assessments of other agencies included in the sample. To the best of our knowledge, we have excluded any source of governance data found to be explicitly based on another one of our sources. Nevertheless, the possibility of correlated errors remains. The main consequence would be that our standard errors will be biased downward, see Kaufmann and others (1999a) for an example. This underscores the importance of caution in comparing governance estimates across countries and over time.

This standard deviation is declining in the number of individual indicators in which a particular country appears and increasing in the variance of the disturbance term on each of these indicators.

The assumptions of the unobserved-components model ensure that the distribution of governance in each country is normal, conditional on the data for that country. Therefore, these conditional means and standard deviations for each country have a natural interpretation: there is a 90 percent probability that the true level of governance in a country is in an interval of plus or minus 1.64 times the reported standard deviation centered on the point estimate itself. This range is referred to as a 90 percent confidence interval around the estimate of governance for a country.[9]

Implementing his approach requires estimates of all the unknown survey-specific parameters, $\alpha(k)$, $\beta(k)$, and $\sigma^2_\varepsilon(k)$. These are computed in a two-stage procedure. In the first stage, maximum likelihood methods are applied, using only the representative sources to retrieve the parameters for each governance cluster. This is a standard application of the unobserved-components model. The many nonrepresentative sources cannot be included in the first stage of the estimation procedure because the distribution of unobserved governance in the subset of countries covered by these surveys is not the same as that in the world as a whole. As a result, for these sources governance in the countries covered by these surveys cannot be assumed to follow a standard normal distribution, as is required by the maximum likelihood procedure.

The parameters of the nonrepresentative sources are obtained in the second stage of the procedure. In this stage the preliminary estimates of governance based only on representative sources are treated as an observable proxy for governance, and the parameters of interest for the nonrepresentative sources are obtained by regressing these indicators on observable governance (directly estimating equation 1).[10] All the estimated parameters of the unobserved-components model are then used to construct a final set of estimates of governance.

The resulting estimates have an expected value across countries of zero and a standard deviation across countries of one. Due to sampling variability, this will not be exactly true for any individual governance indicator in any period. To avoid confusion about the units of the governance indicators, the estimates of governance are rescaled by subtracting the mean across countries and dividing by the standard deviation across countries for each indicator, so that each indicator has a mean of zero and a standard deviation of one in each period.

9. This is a slight abuse of terminology, because these are not confidence intervals in the usual sense of a stochastically varying interval centered around a fixed unknown parameter. Rather, governance is treated as a random variable, and the 90 percent confidence interval is simply the 5th and 95th percentiles of the conditional distribution of governance, given the observed data.

10. Getting consistent estimates of the parameters of the nonrepresentative sources requires adjustment for attenuation bias caused by the fact that the observable proxy for governance is a noisy indicator of true governance. Fortunately, the information on the standard errors associated with the governance estimates obtained in the first stage can be used to do this.

It is also important to note the assumption that the distribution of unobserved governance is the same in every period. In particular, this imposes the restriction that the mean or world average of governance is the same in each period. As a result, the indicators are not informative about global trends in governance (see section V), although they are informative about changes in countries' relative positions over time.

II. Governance Indicators for 1996, 1998, 2000, and 2002

This section presents aggregate governance indicators for the six indicators for all four periods and also examines changes in the indicators over time. The data are available online at www.worldbank.org/wbi/governance/govdata2002/index.html.

Levels of Governance Worldwide

The governance estimates are normally distributed with a mean of zero and a standard deviation of one in each period, implying that virtually all scores lie between −2.5 and 2.5 (higher scores indicate better outcomes). As discussed, this also implies that the aggregate estimates convey no information about global trends in governance. They are, however, informative about changes in countries' relative positions over time.

The voice and accountability indicator covers 199 countries for 2002, the largest set among the six indicators (table 2).[11] Four indicators cover 195 countries, and one, political stability, covers 186 countries. Over time, there has been a steady increase in the number of countries covered by each indicator. The number of data sources has increased as well, raising the median number of sources available per country, which ranges from six to eight in 2002 compared with four to six in 1996. The proportion of countries in the sample with governance estimates based on only one source has declined considerably, from an average of 15 percent in 1996 to 10 percent in 2002. Because the 2002 indicators now cover virtually all countries in the world, no major improvements in country coverage are expected in the future.

An important consequence of this expanded data availability is that the standard errors for the governance indicators have declined. In 1996 the average of the standard errors ranged from 0.26 to 0.39, whereas in 2002 they ranged from 0.19 to 0.27 (see table 2). Moreover, the average standard errors for the revised 1998 and 2000 indicators are also lower than the previous estimates, again reflecting the incorporation of more data for more countries. These

11. A few of the entities covered by the indicators are not independent states (French Guyana; Hong Kong, China; Martinique; Puerto Rico; and West Bank and Gaza). A handful of very small independent principalities (Andorra, Monaco, and San Marino) are also included. For convenience, all 199 entities are referred to as countries.

TABLE 2. Summary Statistics on Governance Indicators

Year	Voice and accountability	Political stability	Government effectiveness	Regulatory quality	Rule of law	Control of corruption	Overall[a]
Number of countries							
1996	192	165	180	182	167	151	173
1998	192	166	184	185	186	184	183
2000	192	166	185	186	186	185	183
2002	199	186	195	195	195	195	194
Median number of sources per country							
1996	4	4	4	4	6	4	4
1998	4	4	4	4	7	5	5
2000	5	6	5	4	8	6	6
2002	7	6	6	6	8	7	7
Proportion of countries with only one data source							
1996	0.15	0.13	0.21	0.15	0.07	0.18	0.15
1998	0.14	0.10	0.19	0.13	0.11	0.18	0.14
2000	0.14	0.06	0.18	0.13	0.10	0.17	0.13
2002	0.10	0.11	0.10	0.10	0.10	0.10	0.10
Average standard error							
1996	0.26	0.39	0.28	0.34	0.26	0.29	0.30
1998	0.25	0.32	0.30	0.34	0.25	0.25	0.28
2000	0.25	0.33	0.27	0.35	0.22	0.26	0.28
2002	0.21	0.27	0.22	0.22	0.19	0.21	0.22

[a]A simple average of the six indicators.

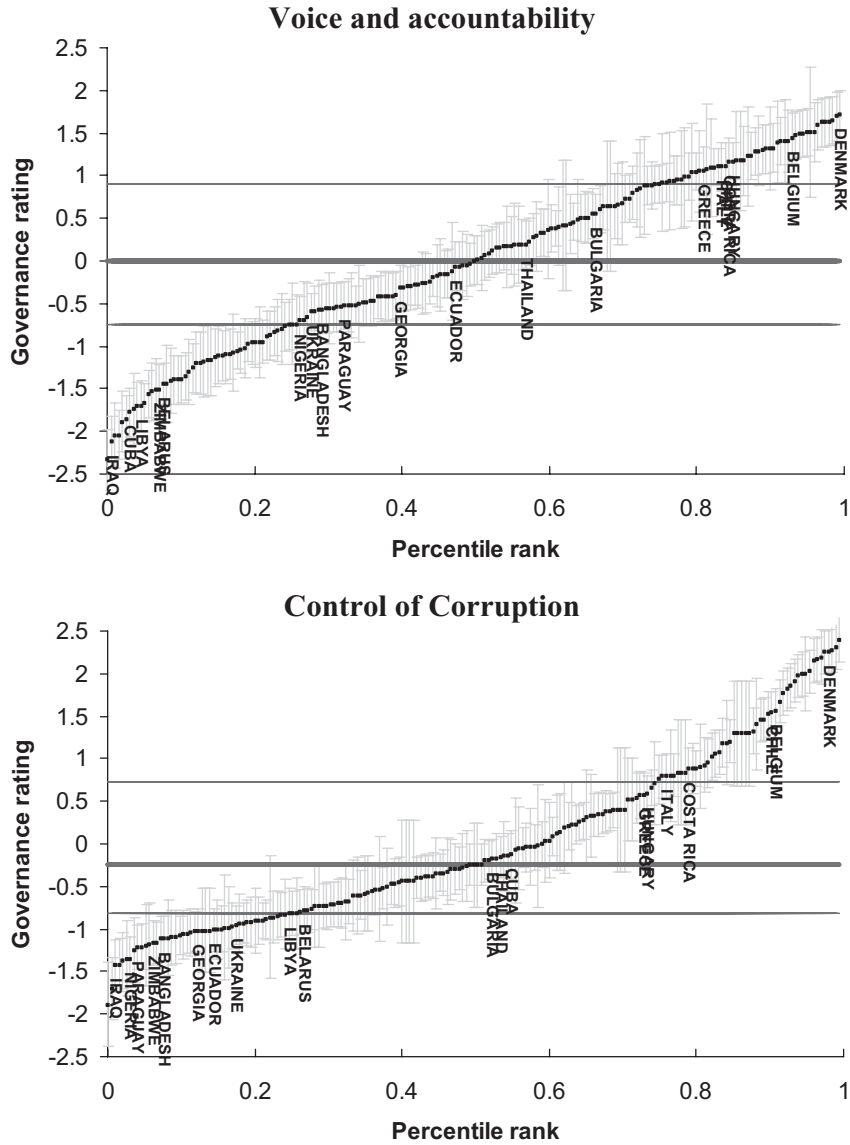
Source: Authors' calculations based on sources listed in table 1.

declines in margins of error illustrate the benefits to precision from constructing composite indicators based on as much information as possible.

Despite these improvements the margins of error associated with estimates of governance remain substantial relative to the units in which governance is measured. This is illustrated by plotting countries in ascending order according to their point estimates of two selected dimensions of governance in 2002 on the x axis and according to estimates of governance and the associated 90 percent confidence interval described above on the y axis for each indicator (figure 1). The confidence intervals vary across countries because countries appear in different numbers of sources with different levels of precision and are large relative to the units in which governance is measured. This point is emphasized by the horizontal lines in figure 1, which delineate the quartiles of the distribution of governance estimates. Even though the differences between countries in the bottom and top quartiles are substantial, the number of countries that have 90 percent confidence intervals that lie entirely within a given quartile is not large.

Thus many of the small differences in estimates of governance across countries are not likely to be statistically significant. For many applications it is therefore more useful to focus on the range of possible governance values for each country (as summarized in the 90 percent confidence intervals shown in figure 1) than on the point estimates. For two countries at opposite ends of the

FIGURE 1. Selected Aggregate Indicators of Governance, 2002



Note: Vertical bars show the statistically likely range of values of governance for each country, with the midpoint of each bar corresponding to the best single estimate. Selected countries are labeled.

Source: Authors' calculations based on sources listed in table 1.

scale of governance whose 90 percent confidence intervals do not overlap at all, it is clear that differences in governance between them are significant. For pairs of countries that are closer together and whose 90 percent confidence intervals overlap, the significance of estimated differences in governance is less clear.

Differences across countries in the margins of error associated with governance estimates are due to cross-country differences in the number of sources per country and differences in the precision of the sources in which each country appears. As equation 3 shows, the precision of governance estimates increases with the number and precision of sources. Across countries the standard error of the governance estimate for a country declines at the rate of the inverse of the square root of the number of sources, consistent with the assumption that errors are uncorrelated across sources. In practice, this means that a 90 percent confidence interval for a country with only one source will be roughly twice as large as the 90 percent confidence interval for a country appearing in the median number of seven sources.

On the precision of sources, recall that for each source the variance of the error with which it measures the unobserved true level of governance is estimated and the inverse of these estimated variances is used to weight sources when constructing the aggregate score for each country. This means that more precise sources (in the sense of providing less noisy signals of governance) receive more weight in the aggregate indicators, minimizing the variance in the estimates of governance for each country. There is considerable variation in the weights assigned to different sources (reflecting substantial differences in the estimated precision of each source),[12] and these differences are reflected in the differences in the margins of error associated with governance scores for each country. Table 3 summarizes the weights applied to each source for a hypothetical country appearing in all sources in constructing the corresponding aggregate indicator. The weights for a country appearing in a subset of sources would be proportional to the ones reported for those sources.

Changes over Time in Estimates of Governance

The observed change in governance for a given country between two points in time can be attributed to four factors: changes in the perceptions of governance recorded in the underlying sources available in both periods, changes in the weights applied to different sources in each period, changes to the set of sources for a country, and the addition of new countries to the aggregate indicator that systematically rate better or worse than the country in question (recall that indicators measure only countries' relative positions).

Changes in scores assigned to countries by underlying sources is the clearest reason for changes in governance for a country over time. Interpreting changes due to the other three reasons is more difficult and involves tradeoffs. Consider the reweighting of sources that occurs from year to year as the observed

12. In estimating the unobserved components model, the estimated precision of each source reflects the extent to which that source is correlated with other sources. In the empirical framework, errors are assumed to be uncorrelated across sources. As a result, sources that tend to be highly correlated with other sources are more informative, and hence have lower error variances, than sources that tend to be only weakly correlated with other sources.

correlations among sources change. The reweighting ensures the most precise estimates of the level of governance for each year, but some of the changes over time for a given country will reflect this reweighting rather than changes in the underlying indicators. However, these changes in weights account for only a small fraction of the variance of changes in governance estimates on average (Kaufmann and others 2002).

There are also tradeoffs in interpreting the changes in governance estimates from the addition of new sources for a country. Adding new data sources improves the precision of estimates of governance in a country at a single point in time. However, if the new sources rate a country significantly differently from existing sources, this can result in changes in estimates of governance that reflect the inclusion of new information on the previous period rather than actual changes in governance. To reduce the effects of this source of variation in governance estimates and to improve the precision of estimates for past years, previous indicators have been recalculated incorporating all the data used for this analysis. Nevertheless, the 2002 indicators also reflect the information embodied in a few new sources relative to 2000 and previous periods, and this provides a further reason why changes over time should be interpreted with caution.[13]

Bias resulting from the addition of new countries to the aggregate indicator can be removed from comparisons of governance estimates over time by limiting such comparisons to changes in countries' percentile ranks for the same set of countries for both periods. This procedure is a useful robustness check when considering changes over time in a specific country or set of countries. In practice, however, this source of bias is relatively small, especially when comparing 2002 with 2000, because there are now only small changes in the number of countries covered between these two periods.

A final issue concerns the statistical significance of observed changes in the aggregate indicators. The basic observation is that changes in the estimates of governance tend to be small relative to the cross-country differences in levels of governance. It is difficult to be more precise about the statistical significance of changes in governance because of the aggregation procedure. For each period the aggregation procedure summarizes knowledge about governance in a given country in terms of the distribution of unobserved governance conditional on the data for that country. The mean of this conditional distribution is used as the best estimate of the level of governance in a country, and the standard deviation of this distribution is used to summarize the precision of the knowledge about governance for that country. However, when the aggregation procedure is repeated in successive periods, no information is produced about the joint distribution of governance in successive periods. Without this joint

13. On the Web site displaying the data (www.worldbank.org/wbi/governance/govdata2002/index.html), users may identify the sources of governance data used for each country, indicator, and period.

TABLE 3. Weights Used to Aggregate Governance Indicators

	Voice and accountability				Political stability				Government effectiveness			
	1996	1998	2000	2002	1996	1998	2000	2002	1996	1998	2000	2002
Representative sources												
CDU	—	—	0.03	0.04	—	—	0.07	0.09	—	—	0.07	0.05
DRI	—	—	—	—	0.25	0.17	0.19	0.14	0.08	0.07	0.12	0.07
EIU	0.31	0.12	0.09	0.18	0.13	0.14	0.23	0.21	0.19	0.18	0.20	0.21
FRH	0.14	0.22	0.14	0.12	—	—	—	—	—	—	—	—
HER	—	—	—	—	—	—	—	—	—	—	—	—
HUM	0.05	0.10	0.06	0.04	0.10	0.10	0.08	0.04	—	—	—	—
PRS	0.09	0.14	0.14	0.07	0.09	0.25	0.07	0.07	0.05	0.02	0.07	0.05
RSF	—	—	—	0.02	—	—	—	—	—	—	—	—
WBS	—	0.01	0.00	—	—	0.07	0.01	—	—	0.06	0.03	—
WMO	—	—	—	0.06	—	—	—	0.17	—	—	—	0.13
Nonrepresentative sources												
AFR	—	—	—	0.01	—	—	—	—	—	—	—	0.01
BPS	—	—	—	—	—	—	—	—	—	—	—	0.01
BRI	—	—	—	—	0.22	0.10	0.12	0.11	0.09	0.06	0.08	0.06
EBR	—	—	—	—	—	—	—	—	—	—	—	—
FHT	0.30	0.37	0.49	0.39	—	—	—	—	0.31	0.15	0.12	0.24
GAL	—	—	—	0.01	—	—	—	—	—	—	—	—
GCS	—	—	—	0.02	0.13	0.07	0.04	0.05	0.06	0.09	0.09	0.06
GCSA	—	—	—	—	—	0.04	—	—	—	0.15	—	—
GMS	—	—	0.02	—	—	—	0.01	—	—	—	0.00	—
LOB	0.08	—	—	0.01	—	—	—	0.03	0.00	—	—	0.01
OPF	—	—	—	—	—	—	—	—	—	—	—	—
PIA	—	—	—	—	—	—	—	—	0.13	0.15	0.12	0.06
QLM	—	—	—	—	—	—	—	—	—	—	—	—
WCY	0.01	0.00	0.01	0.01	—	—	0.13	0.06	0.05	0.04	0.06	0.04

— not available.

Note: See table 1 for source codes. The weights used in constructing the aggregate governance indicators correspond to those that would be applied for a hypothetical country appearing in all of the available sources for that indicator. The weights are proportional to the inverse of the variance of the estimate of measurement error for each source (see discussion in text). For a country appearing in fewer sources, the relative weights applied to each source will be the same as the relative weights implicit in this table.

Source: Authors' calculations based on sources listed in table 1.

distribution, precise probabilistic statements cannot be made about changes in governance over time.[14]

Instead, a useful rule of thumb is to focus on changes in governance for countries in which the 90 percent confidence intervals in the two periods do not overlap. This can be illustrated by plotting the 2002 score on the x axis and the

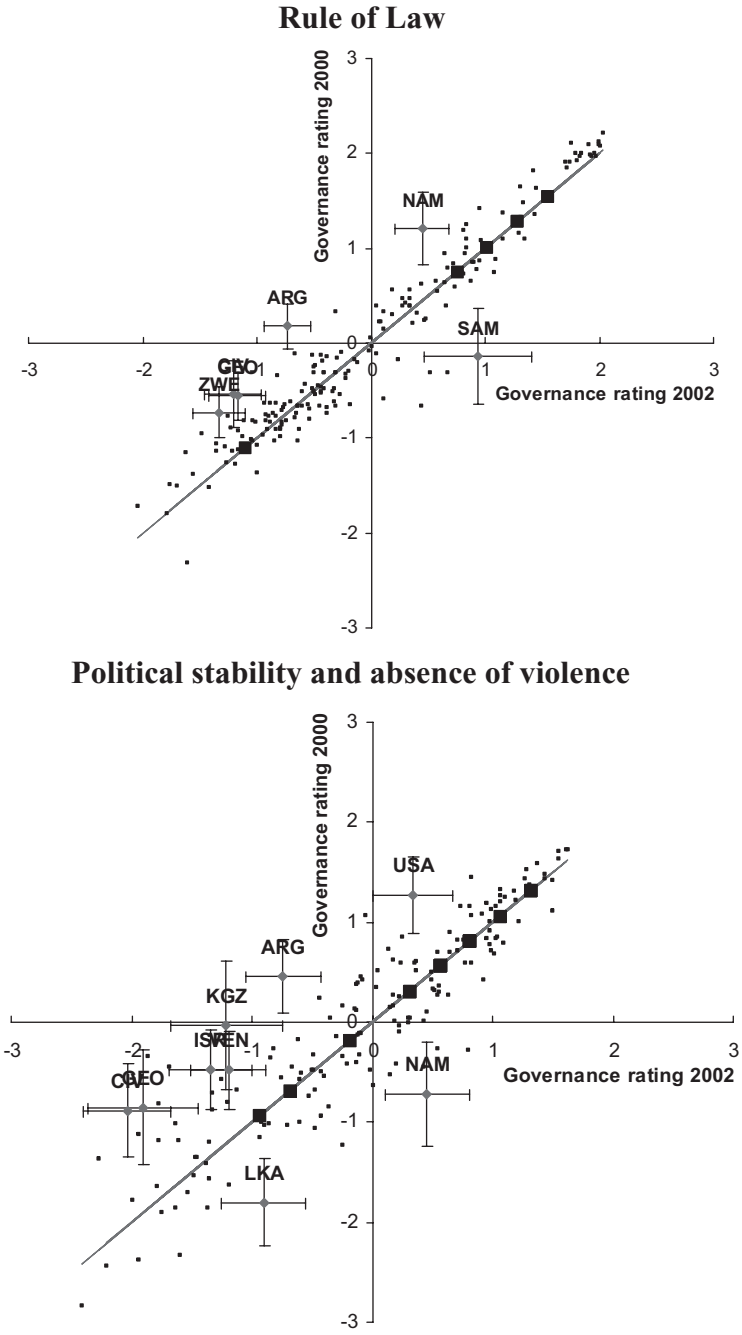
14. Extensions of the aggregation procedure along the lines of dynamic unobserved component models could in principle provide information about the joint distribution of governance over time. We have not yet attempted to implement this idea with our data.

Regulatory quality				Rule of law				Control of corruption			
1996	1998	2000	2002	1996	1998	2000	2002	1996	1998	2000	2002
—	—	—	—	—	—	0.03	0.03	—	—	0.08	0.06
0.09	0.08	0.23	0.03	0.03	0.06	0.08	0.09	0.09	0.05	0.09	0.06
0.23	—	—	0.21	0.34	0.08	0.06	0.20	0.24	0.07	0.11	0.12
—	—	—	—	—	—	—	—	—	—	—	—
0.10	0.05	0.09	0.06	0.05	0.08	0.07	0.08	—	—	—	—
—	—	—	—	0.02	0.02	0.02	0.01	—	—	—	—
0.03	0.03	0.06	0.09	0.03	0.04	0.03	0.02	0.05	0.01	0.03	0.03
—	—	—	—	—	—	—	—	—	—	—	—
—	0.00	0.01	—	—	0.05	0.02	—	—	0.05	0.07	—
—	—	—	0.26	—	—	—	0.11	—	—	—	0.09
—	—	—	—	—	—	—	—	—	—	—	0.02
—	—	—	0.00	—	—	—	0.00	—	—	0.08	0.01
—	—	—	—	0.08	0.06	0.09	0.06	0.03	0.01	0.01	0.01
0.07	0.23	0.06	0.11	—	—	—	—	—	—	—	—
—	—	—	—	0.16	0.17	0.27	0.13	—	0.18	0.23	0.22
—	—	—	—	—	—	—	0.01	—	—	—	—
0.14	0.19	0.09	0.04	0.07	0.10	0.14	0.07	0.09	0.12	0.10	0.06
—	0.15	—	—	—	0.00	—	—	—	0.17	—	—
—	—	—	—	—	—	0.00	—	—	0.04	0.02	—
—	—	—	—	0.00	—	0.00	—	—	—	—	0.06
—	—	—	—	—	—	—	—	—	—	—	—
0.08	0.08	0.21	0.07	—	0.11	0.10	0.04	—	0.11	0.08	0.04
—	—	—	—	0.11	0.12	—	0.07	0.10	0.10	—	0.11
0.21	0.16	0.20	0.10	0.10	0.08	0.08	0.06	0.36	0.07	0.09	0.08

2000 score on the *y* axis and by drawing the 45° line that distinguishes countries with declines in the quality of governance from countries with improvements in governance (figure 2). Countries with large changes in governance relative to their margins of error in each period are highlighted, and the 90 percent confidence intervals in each period is indicated by vertical and horizontal lines. The score for each country that appears in the 2002 indicators but not in the 2000 indicators is plotted along the 45° line, giving a visual summary of the distribution of governance among the countries added to the sample in 2002.

The number of countries with large changes in governance over this brief period is quite small. This is not surprising given the short period under consideration and the gradual nature of most changes in governance. Many of

FIGURE 2. Changes Over Time in Selected Governance Estimates, 2000 to 2002



Note: Highlighted countries are those in which the 90 percent confidence intervals in the two periods do not overlap. The corresponding confidence intervals in 2000 and 2002 are indicated as vertical and horizontal bars. The 45° line demarcates the difference between countries showing declines in governance (above the line) and those showing improvements (below the line).

Source: Authors' calculations based on sources listed in table 1.

the changes are understandable in light of events over these two years. For example, Argentina's recent financial crisis is reflected in strong declines in perceptions of governance across the board. Similarly, the recent turmoil in Zimbabwe is associated with a sharp decline in perceptions of the rule of law. For the United States, declines in political stability and absence of violence reflect heightened concerns about terrorism in the aftermath of the terrorist attacks of September 11, 2001. In Sri Lanka reductions in sectarian violence drive an improved score in this category.

The reasons for some of the other changes highlighted in figure 2 are less obvious. These are examined in more detail by tallying the number of sources available in both periods that move in the same or opposite direction as the aggregate indicator or that register no change (table 4). The overall rate of agreement between changes in the sources and the direction of change in the aggregate indicator is calculated as the agreement ratio (the ratio of number of agreements to the total number of changes in both directions).

The agreement ratio is quite high for countries with large changes in governance, with an average across all countries and indicators of 0.79. This provides some confidence that for countries with large changes in governance estimates, the changes are driven primarily by changes in underlying sources. For only four countries is the agreement ratio less than one-half—Belarus and Iraq for regulatory quality and Madagascar and West Bank and Gaza for control of corruption. Belarus's surprisingly high score in 2002 is driven primarily by the very strong responses from firms in the Business Environment and Enterprise Performance Survey. Iraq illustrates an unusual case in which the reweighting of sources has a substantial effect on changes over time. In both periods Iraq has one of the worst scores in the world, so the large change in its score reflects no real improvement during the period but rather the much lower weight assigned to the source that rated Iraq highest in 2002. For control of corruption the large improvement observed in Madagascar and the large decline in West Bank and Gaza are both driven entirely by changes in the set of sources in which these countries appear. Madagascar appears in one new source that rates it highly (World Markets Online), and it does not appear in the 2002 version of the State Capacity Project, which gave it a poor score in 2000. West Bank and Gaza fares well on the World Business Environment Survey in 2000, a source that is not available in 2002, and scores poorly on the only source available for 2002, World Markets Online.

Unlike these large changes in governance, which reflect primarily changes in the underlying sources, the majority of smaller changes reflect a combination of all four reasons for variation discussed. For the remaining smaller changes in governance between 2000 and 2002 not reported in table 4, the average agreement ratio across all countries ranges from 0.57 to 0.64 for the six indicators, substantially lower than the agreement ratio for large changes. This suggests a need for greater caution in interpreting the small changes in governance estimates typical from one period to the next. Although changes over longer

TABLE 4. Large Changes in Governance Estimates, 2000 to 2002

Indicators and countries	Governance score		Sources available in both periods[a]				Changes in sources between 2000 and 2002	
	2002	2000	Agree	No change	Disagree	ratio	Added	Dropped
Voice and accountability								
Sierra Leone	−0.57	−1.36	2	2	0	1.00	2	0
Political stability								
Sri Lanka	−0.90	−1.80	3	0	3	0.50	1	0
Namibia	0.46	−0.72	3	0	1	0.75	3	1
Argentina	−0.74	0.46	7	0	1	0.88	2	1
Côte d'lvoire	−2.04	−0.88	3	0	1	0.75	2	1
Georgia	−1.90	−0.85	3	0	0	1.00	1	2
Israel	−1.35	−0.47	4	0	3	0.57	1	0
Kyrgyz Republic	−1.21	−0.03	1	1	0	1.00	1	1
United States	0.34	1.26	5	1	1	0.83	1	2
Venezuela	−1.20	−0.48	4	1	2	0.67	3	1
Government effectiveness								
Dominica	0.32	−0.86	1	0	0	1.00	1	0
Argentina	−0.49	0.30	7	1	0	1.00	1	1
Egypt, Arab Rep.	−0.32	0.35	4	2	1	0.80	1	1
Gambia, The	−0.81	0.25	1	0	1	0.50	2	0
Tunisia	0.65	1.32	4	0	1	0.80	1	1
Regulatory quality								
Afghanistan	−1.82	−3.57	1	0	0	1.00	1	0
Belarus	−1.67	−2.65	1	1	2	0.33	3	1
Iraq	−2.31	−3.36	0	0	3	0.00	2	0
Moldova	0.80	0.14	5	0	0	1.00	3	1
Russian Federation	−0.30	−1.55	6	0	1	0.86	3	1
Congo, Dem. Rep.	−1.77	−2.87	3	0	0	1.00	1	1
Argentina	−0.84	0.44	5	0	1	0.83	2	1
Bangladesh	−1.05	−0.02	2	2	0	1.00	3	1
Cameroon	−0.88	0.12	3	1	0	1.00	2	1
EI Salvador	0.04	1.12	2	1	1	0.67	2	1
Zambia	−0.60	0.43	3	0	1	0.75	2	1
Rule of law								
Samoa	0.94	−0.14	1	0	0	1.00	1	0
Argentina	−0.73	0.18	8	1	2	0.80	2	4
Côte d'lvoire	−1.21	−0.53	2	3	1	0.67	2	1
Georgia	−1.17	−0.56	4	3	0	1.00	2	2
Namibia	0.45	1.21	2	3	1	0.67	2	1
Zimbabwe	−1.33	−0.73	6	2	0	1.00	1	0
Control of corruption								
Madagascar	0.14	−0.80	0	1	1	0.00	1	2
Belarus	−0.78	−0.07	4	1	0	1.00	1	2
Malawi	−0.91	−0.22	2	2	0	1.00	2	2
Namibia	0.21	1.16	3	0	1	0.75	4	1
West Bank and Gaza	−0.99	0.76	0	0	0	—	1	1
Average			3.11	0.78	0.81	0.79	1.78	1.00

— not available.

Note: Sources of changes in estimates of governance between 2000 and 2002 for each country for which the 90 percent confidence intervals for the level of governance in the two periods do not overlap.

[a]The number of individual sources that agree or disagree with the direction of change of the aggregate indicator.

Source: Authors' calculations based on sources listed in table 1.

periods of time, such as from 1996 to 2002, would be expected to be more informative, this is partly offset by the larger changes in the composition of the sources between the periods.

III. Use of Perceptions-Based Sources and Potential Ideological Bias

The construction of the aggregate governance indicators described herein relies exclusively on subjective, perceptions-based measures of governance. For many of the key dimensions of governance, such as corruption or the confidence that property rights are protected, objective data are almost by definition impossible to obtain, so there are few alternatives to subjective data.

Why Perceptions-Based Measures Are Used

Consider corruption. As an illegal activity, there are no direct measures of its prevalence. Various indirect measures are possible, but none are without difficulty. For example, relying on the frequency of references to corruption in the media will reflect not only the prevalence of corruption but also the freedom and objectivity of the press. Similarly, trials for corruption will reflect the competence and independence of the police and the judicial system and not exclusively the prevalence of corruption. Recently, a few studies have attempted to assess corruption by looking for patterns in objective data that can only be consistent with corruption, such as variations in the procurement prices paid for homogeneous medical inputs across hospitals in Buenos Aires (Di Tella and Shargrodsky 2003) and gaps between existing stocks of public infrastructure and past infrastructure spending across Italy (Golden and Picci 2003). Though interesting, such exercises have enormous data requirements, and cross-country measures of corruption based on this idea are unavailable.[15]

Objective measures may be available for some other dimensions of governance, but they are not without weaknesses. Objective data on elections can be used to measure democratic participation. But there is considerable variation across countries in the extent to which election outcomes reflect the will of voters. Measuring the extent to which elections are subverted, whether through intimidation or manipulation of results, returns quickly to the realm of perceptions-based data. This is just one example of the important distinction between de jure and de facto situations regarding governance. Countries may have extensive formal protections of property rights but little or no enforcement.

15. Furthermore, these within-country measures based on prices, assets, and expenditure patterns are typically a proxy of the combined effect of the extent of mismanagement, inefficiency, and corruption. Disentangling the pure effect of corruption is far from simple.

For example, most countries now have formal independent anticorruption commissions or agencies, but their effectiveness varies greatly.

Subjective perceptions of governance often matter as much as the legal reality. For example, Hellman and Kaufmann (2004) develop a measure of "crony bias" or unequal influence across firms based on firms' perceptions of undue influence on political decisionmaking exerted by powerful firms. They find that perceived unequal influence has a strong negative impact on a firm's assessment of public institutions and on its behavior toward those institutions, resulting in less use of the courts to resolve business disputes, lower enforceability of court decisions, lower levels of tax compliance, and higher levels of bribery. Thus, inequality of influence not only damages the credibility of institutions but also affects the likelihood that firms will use and provide tax resources to support such institutions, thereby perpetuating the weakness of such institutions and the likelihood of capture by the influential.

Finally, recent studies have yielded a profusion of results linking objective measures of the structure of institutions to a range of governance outcomes.[16] Although the studies have greatly expanded understanding of the institutional determinants of development, these objective measures of institutional quality do not lend themselves well to the construction of aggregate governance indicators. The measures typically do not have normative content on their own. They assume normative content only in the context of a particular empirical analysis linking them with a particular outcome. For example, although measures of decentralization may be correlated with the incidence of corruption across countries, the explanatory power of this variable is generally not sufficiently strong to consider decentralization as a reasonable proxy for corruption.

None of this is to suggest that the subjective data used here are problem-free. Perceptions-based questions about governance can be vague and open to interpretation. For example, a well-crafted question on corruption asks firms for the estimated share of bribes in the annual spending of firms like theirs. By contrast, generalized opinion questions, such as a citizen's perception of the overall tolerance of the population for corruption, are less informative for constructing aggregate indicators of governance.

Today, studies like this can rely on more specific, better crafted, and to an extent experiential questions. For instance, the Global Competitiveness Report survey of firms contains much more specific questions about corruption and governance than even during the mid-1990s, and some are of a quantitative and experiential nature (such as percentage of senior management time spent dealing with public officials).

16. A nonexhaustive list includes the links between decentralization and corruption, the effect of the structure of the legal system on financial market development, the effect of checks and balances in the political system on regulatory and fiscal performance, and the effect of democratic institutions on a wide variety of development outcomes.

Potential Ideological Biases in Perceptions Data

A potential drawback of information collected from polls of experts is that it may reflect the ideological tendencies of the institutions compiling the performance ratings. This may not be a major concern for the sources used for this analysis. The high degree of correlation among virtually all of the sources is difficult to reconcile with a systematic ideological bias among one or more sources. Nevertheless, it is useful to investigate the extent to which differences in assessments are related to observable measures of the ideology of the government in power in each country.

This is done as follows. Surveys of firms or individuals are assumed not to be tainted by ideology, because they reflect the views of a large number of respondents in each country. However, a poll of a smaller number of experts affiliated with a particular institution may reflect that institution's ideology. The effects of ideology can therefore be identified by looking at correlation across countries of the ideology of the government in power and the difference in the percentile ranks assigned to countries by a poll of experts and a survey of individuals and firms. This approach was applied to several polls of experts, using the World Business Environment Survey for 2000 as a benchmark survey assumed to be unaffected by respondent ideology. Government ideology was measured using an indicator variable of the political orientation of the government in power (taken from Beck and others 2001) that takes on the value 1 for left of center, 2 for centrist, and 3 for right of center.

The difference between the percentile rank of a country on a poll of experts and its rank on the World Business Environment Survey is regressed on the indicator variable measuring the ideology of the government in that country, for several polls of experts (table 5). All variables are measured in 2000, the most recent year for which the ideology variable is available.[17] The coefficient on the ideology variable will therefore capture the extent to which a given poll of experts rates countries with left- or right-wing governments systematically differently from the survey. A positive coefficient indicates that the poll tends to rate right-of-center governments more highly relative to the survey, whereas a negative coefficient indicates a bias toward left-of-center governments.[18] The Heritage Foundation is the only source that appears to have a consistent ideological bias, assigning relatively higher scores to countries with right-of-center governments than the corresponding surveys. However, this ideology bias is fairly modest, resulting in about a 7–10 percentage point higher ranking for a right-of-center government

17. For voice and accountability, Gallup Millennium Survey is used instead of World Business Environment Survey as the comparator survey, because the World Business Environment Survey questions on voice and accountability capture the extent to which firms have a voice in policymaking, which is considerably narrower than most other polls.

18. Because most of the countries are coded as left or right of center, almost identical results are obtained if dummy variables for left- and right-of-center governments are included instead.

TABLE 5. Ideology Regressions for 2000

	PRS	PIA	EIU	DRI	CDU	BRI	QLM	HUM	EBR	HER	FRH
Voice and accountability[a]											
Ideology	−2.78		−1.64		−1.72			3.67			−0.83
	0.59		0.46		0.27			0.68			0.23
Observations (no.)	44		43		28			46			46
Adjusted R^2	−0.01		−0.02		−0.04			−0.01			−0.02
Political stability											
Ideology	12.37		8.86	8.54	4.97	3.15		12.11			
	2.68**		1.80*	1.87*	0.93	0.61		2.52**			
Observations (no.)	52		51	46	42	25		56			
Adjusted R^2	0.1		0.04	0.05	−0.01	−0.02		0.09			
Government effectiveness											
Ideology	−1.84	−0.66	−2.38	1.86	−7.12	1.64					
	0.64	0.16	0.68	0.48	1.90*	0.25					
Observations (no.)	52	47	51	46	42	25					
Adjusted R^2	−0.01	−0.02	−0.01	−0.02	0.05	−0.04					
Regulatory quality											
Ideology	8.05	13.3		3.22					6.55	10.24	
	1.57	2.08**		0.45					0.88	1.77*	
Observations (no.)	52	47		46					15	56	
Adjusted R^2	0.02	0.07		−0.02					−0.01	0.04	
Rule of law											
Ideology	1.52	3.39	5.61	5.67	4.68	7.32	6.47	5.32		7.42	
	0.41	0.73	1.65	1.46	1.21	1.65	1.63	1.19		1.91*	
Observations (no.)	52	47	51	46	42	25	49	56		56	
Adjusted R^2	−0.02	−0.01	0.03	0.02	0.01	0.05	0.03	0.01		0.05	
Control of corruption											
Ideology	3.05	1.4	0.31	0.57	−2.21	2.83	1.84				
	0.63	0.34	0.1	0.18	0.68	0.46	0.58				
Observations (no.)	52	47	51	46	42	25	49				
Adjusted R^2	−0.01	−0.02	−0.02	−0.02	−0.01	−0.03	−0.01				

*Significant at the 10 percent level.

**Significant at the 5 percent level.

Note: See table 1 for source codes. Results of cross-country regressions of difference in percentile rank between each poll of experts and the corresponding question from the World Business Environment Survey on an indicator variable taking the value 1 if the government of a country is left of center, 2 if it is center, and 3 if it is right of center. Percentile ranks are on a scale from 0 to 100, based on the sample of countries common to each pair of sources. The table reports the slope coefficient and *t*-statistic.

[a]Uses a question from the Gallup Millennium Survey instead of World Business Environment Survey.

Source: Authors' calculations based on sources listed in table 1.

than for a center government. Moreover, in all cases the ideology variable explains only a trivial fraction of the difference in assessments between polls and surveys, suggesting that the importance of ideological biases in polls is quite small overall.

IV. Margins of Error and Classifying Countries According to Governance Performance

Margins of error are not unique to subjective indicators but are pervasive in all efforts to measure governance. The margins of error complicate the use of governance indicators for classifying countries according to governance performance. Classifications based on individual indicators or even on a single aggregate indicator inevitably run the risk of misclassifying countries due to the margins of error inherent in all indicators.

Margins of Error Are Not Unique to Subjective Data

One of the strengths of the governance indicators reported here is the ability to construct explicit margins of error for the estimates of governance for each country. These margins of error are not unique to subjective or perceptions-based measures of governance, however, but apply to most other measures of institutional quality and to many other socioeconomic indicators as well. That measurement error is pervasive is obvious from the range of "preliminary" estimates of basic variables such as real GDP growth produced even in countries with high-quality statistical systems.

Consider, for example, recent efforts to construct measures of governance that rely on objective and quantifiable data rather than exclusively on perceptions-based data sources. Knack and Kugler (2002) argue that variables such as waiting time to obtain a telephone line and number of telephone faults can serve as proxies for public administrative capacity, that degree of government reliance on trade taxes can serve as a proxy for the ability of the government to broaden its tax base, or that volatility in budgetary expenditure and revenue shares are indicative of a volatile policy environment. Clague and others (1999) argue that the proportion of currency in circulation held in the banking system is a good proxy for protection of property rights. Djankov and others (2002, 2003) use cross-country data on the number of administrative procedures required to start a business and the number of legal procedures required to collect an unpaid debt to capture the complexity of the regulatory and legal environment.

Although such measures can in principle provide an accurate measure of the underlying concept they attempt to assess, their usefulness as a measure of broader notions of governance depends on how well the underlying concept corresponds to such broader ideas of governance. For example, the number of procedures required to start a business may not be a good indicator of the complexity or burden of regulation in other areas. Similarly, the willingness of individuals to hold currency in banks reflects confidence in a very particular set

of property rights but does not necessarily capture other dimensions of property rights protection, such as confidence in the police and judicial system.

This is not a specific drawback of such objective measures. All measures, subjective and objective, are necessarily imperfect proxies for broader notions of governance. But it does reinforce the importance of margins of error for objective indicators as well, to the extent that they are used as proxies for broad concepts of governance such as those measured here using subjective data.[19] A simple calculation can give a sense of the order of magnitude of margins of error for objective indicators. Suppose that there are two noisy indicators y on a common unobserved concept of governance, $g: y_i = g + \varepsilon_i, i = 1,2$. If the variance of the unobserved measure of governance is normalized to be one, the correlation between the two observed indicators is $\rho = [(1 + \sigma_1{}^2)(1 + \sigma_2{}^2)]^{-1/2}$. Suppose that indicator 1 is one of the subjective governance indicators presented here, for which the variance of the measurement error, $\sigma_1{}^2$, is known, and that indicator 2 is one of the objective indicators already described. Then from the observed correlation between the two indicators, the variance of measurement error in the objective indicator, $\sigma_2{}^2$, can be inferred.

This calculation is done for several objective governance indicators (table 6). The implied standard deviation of measurement error in the objective indicator is calculated under three assumptions: that the estimate of the standard deviation of measurement error in the subjective indicator is correct (assumption A), that the subjective and objective indicators have the same standard deviation of measurement error (assumption B), and that the standard deviation of measurement error in the subjective indicator is twice as large as that in the objective indicator (assumption C). The actual standard deviation of measurement error for the subjective indicator is also calculated, computed as the average across all countries of the country-specific standard errors in our governance indicators.

For all indicators and for all three sets of assumptions, the implied standard deviation of measurement error in the objective indicators is much higher than the corresponding standard deviation of the subjective governance indicator. Under benchmark assumption A, the implied margin of error for the objective indicators is 7–15 times larger than that of the subjective indicators. This clearly exaggerates the difference in the precision of subjective and objective indicators because it compares a single objective indicator with an aggregate of several subjective measures, and as discussed, aggregation improves precision.

But this is only part of the story. The government effectiveness and regulatory quality indicators have a median of six sources per country, and the rule of law indicator has a median of eight sources. This can explain why the standard deviation of measurement error of the objective sources might be $\sqrt{6} = 2.4$ to $\sqrt{8} = 2.8$ times higher than that of the corresponding subjective indicators, but it still cannot explain

19. These margins of error should, of course, also reflect measurement error in the raw data on which they are based—for example, the nontrivial measurement error in macroeconomic variables such as the money supply or the composition of public expenditures.

TABLE 6. Imputed Margins of Error for Objective Governance Indicators

Objective indicator	Corresponding subjective indicator for 2002	Absolute value of correlation	Implied margin of error for objective indicator[a]			Actual margin of error for subjective indicator
			Assumption A	Assumption B	Assumption C	
Telephone wait time	Governance effectiveness	0.56	1.43	0.88	0.58	0.21
Phone faults	Governance effectiveness	0.32	2.92	1.47	1.00	0.21
Trade tax revenue	Governance effectiveness	0.50	1.68	1.00	0.67	0.21
Budgetary volatility	Governance effectiveness	0.50	1.68	1.00	0.67	0.21
Revenue source volatility	Governance effectiveness	0.49	1.71	1.01	0.67	0.21
Contract intensive money	Rule of law	0.57	1.39	0.86	0.57	0.19
Contract enforcement	Rule of law	0.40	2.25	1.22	0.82	0.19
Regulation of entry	Regulatory quality	0.50	1.67	1.00	0.66	0.22
Aggregate objective indicator	Governance effectiveness	0.73	0.88	0.60	0.39	0.21

[a]Assumption A: the estimate of the standard deviation of measurement error in the subjective indicator is correct. Assumption B: the subjective and objective indicators have the same standard deviation of measurement error. Assumption C: the standard deviation of measurement error in the subjective indicator is twice as large as that in the objective indicator.

Source: Authors' calculations based on sources listed in table 1.

all of the difference in the precision of the indicators. Even with an aggregated objective indicator for government effectiveness, the implied standard deviation of measurement error is still four times as large for the objective indicator as for the subjective one (last row of table 6), though the benefits of aggregation should be roughly comparable for the two indicators, with a median of five sources per country for the objective indicator and six for the subjective indicator.

Assumptions B and C are designed to be more favorable to the precision of the objective indicators. Assumption B discards the information in the margins of error that were constructed for the subjective indicator and simply makes the neutral assumption that the subjective and the objective indicators have the same standard deviation of measurement error. This reduces the implied standard deviation of measurement error for the objective indicator relative to the benchmark assumption A, but it remains large at 0.6 for the composite objective indicator and higher for the individual indicators. Assumption C weights things even further in favor of the objective indicators, by assuming that the objective indicator is twice as precise as the subjective indicator. Yet substantial estimates of the standard deviation of measurement error remain, on the order of 0.4 and higher for individual objective indicators.

This simple calculation underscores and helps quantify the intuitive notion that all governance indicators, not just the subjective ones constructed here, are subject to nontrivial margins of error. Care should be taken in making governance comparisons based on any such measures, and wherever possible it is desirable to construct explicit margins of error to aid in these comparisons.

Margins of Error and the U.S. Millennium Challenge Account

To illustrate the importance of taking margins of error into account when classifying countries by level of governance, this section examines the criteria for country eligibility for the new aid program of the U.S. government, the Millennium Challenge Account (MCA). The MCA allocates funds to countries that "govern justly," "invest in people," and "promote economic freedom," placing governance issues center stage (U.S. Department of State 2002).[20] The allocation criteria draw heavily on a number of cross-country measures of the quality of governance, including five of the six governance indicators presented here (all but the political stability indicator). The first round of countries eligible for MCA funds was selected using these criteria in 2004.

The allocation criteria are designed to ensure that funds go to low-income countries with relatively sound policies and institutions. The process starts with the 74 countries that are eligible for concessional lending from the International

20. Details on the MCA can be found online at www.mca.gov. See Radelet (2003) for a detailed discussion of the MCA. The MCA is not the only example of explicit use of governance indicators. For example, the World Bank uses its internal assessments of countries' policy performance, the Country Policy and Institutional Assessment, to allocate concessional lending from its International Development Association.

Development Association that have per capita incomes of less than $1,435 in 2001.[21] This set of countries is rated according to 16 performance criteria covering three dimensions of performance: governing justly (6 criteria), investing in people (4 criteria), and promoting economic freedom (6 criteria). Four of the governance indicators constructed here (voice and accountability, government effectiveness, rule of law, and control of corruption) are performance indicators under the "governing justly" performance dimension; a fifth governance indicator, regulatory quality, is included under promoting economic freedom.

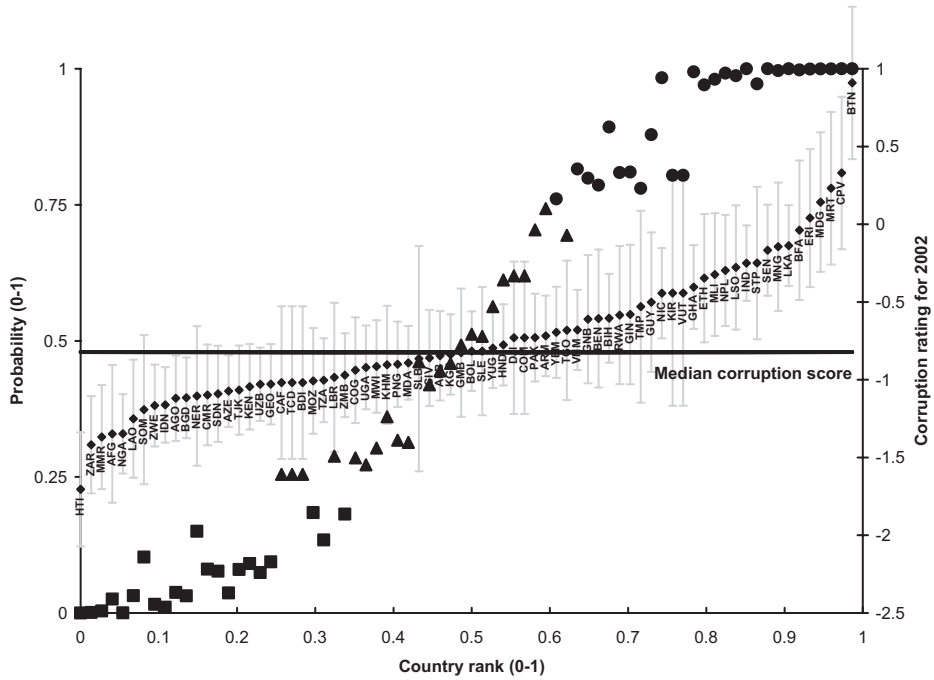
To qualify for assistance, countries must be in the top half of all potentially eligible countries on at least half of the performance criteria under each of the three dimensions of performance. Countries must also be in the top half of potentially eligible countries according to the corruption rating from the governance indicators described in this article. This rule is designed to ensure that resources are channeled to countries that are performing well in a variety of dimensions of governance and in which corruption especially is relatively low.

Though an objective and monitorable set of criteria for determining eligibility is highly desirable, both for allocating aid and for creating clear incentives for potential aid recipients, the substantial margins of error associated with governance estimates mean that it is difficult to assign countries to a definitive performance category based on their estimated level of governance. Recognizing this, the MCA criteria do not require countries to pass the median hurdle on all indicators. However, the allocation rule requires countries to score in the top half of all relevant countries on the corruption indicator, which would constitute a hard hurdle for eligibility. Corruption is surely an important factor in allocating aid, but a simple in-or-out rule runs the risk of misclassifying some countries because of the large margins of error. The MCA fact sheet recognizes this possibility and provides some flexibility for softening this hard hurdle (U.S. Department of State 2002).

To get a sense of the risk of misclassifying countries using a single measure such as corruption, the 74 potential MCA countries are ranked according to their scores on the 2002 control of corruption indicator developed here. This ranking is plotted on the x axis (marked by diamonds), and the 90 percent confidence intervals for each country are shown as a vertical line on the right hand y axis (figure 3). For a substantial fraction of countries, the median score (indicated as a heavy black horizontal line on the graph) falls within the 90 percent confidence interval. For only 11 of the 37 countries in the bottom half of the sample

21. A number of countries with per capita income grater than $1,435 are currently eligible for IDA lending under the small island economies exception, but these will not be eligible for the MCA during the first year. The group of 74 countries is based on data on ida eligibility, available online at www.worldbank.org/ida and per capita gross national income in U.S. dollars in 2001 using *World Bank Atlas* exchange rates, available in the World Bank's *World Development Indicators*. In the second year eligibility will expand to all countries with per capita incomes of less than $1,435, and in the third year to all countries with per capita incomes of less than $2,975.

FIGURE 3. Using Governance Indicators to Allocate Aid for the Millennium Challenge Account



Note: All 74 countries eligible for the first round of the MCA are shown by their corruption rank (on the horizontal axis) and their corruption scores (on the right-hand vertical axis). Diamonds indicate the corruption score, and the vertical lines for each country indicate the 90 percent confidence interval for corruption. The squares, triangles, and circles indicate the probability (on the left axis) that a country's corruption score is in the top half of the sample. Squares indicate countries where this probability is less than 25 percent, circles where it is between 25 percent and 75 percent, and triangles where it is greater than 75 percent.

Source: Authors' calculations based on sources listed in table 1.

is the 90 percent confidence interval fully below the median score. For the remaining 26 countries for which the confidence intervals include the median, there is at least a 5 percent probability that their scores are actually in the top half of the sample. Similarly, only 17 of the 37 countries in the top half of the sample have confidence intervals that are fully above the median score, whereas for the remaining 20 countries there is at least a 5 percent probability that their scores are actually in the bottom half of the sample.

Thus, for the majority of the 74 countries there is a nontrivial probability that they could be mistakenly classified in the wrong half of the sample based on their point estimates of governance alone. The probability that a country's true unobserved level of governance falls in the top half of the sample is plotted on figure 3. Not surprisingly, for the worst rated countries the probability of falling into the top half of the sample is close to zero (marked by squares). Similarly,

the best rated countries almost certainly belong in the top half (the circles). However, for a large intermediate range of 20 countries there is a nontrivial probability that they belong in either the top or bottom half of the sample. For these countries it seems prudent to rely on additional sources of information in making MCA eligibility decisions. This also underscores the need for some flexibility in the MCA allocation rule and the need for this flexibility to be symmetric. Not only should countries that barely miss the list of better performers be given special consideration, as currently proposed in the MCA fact sheet, but countries that barely make the list of better performers also merit further scrutiny.

V. Global Trends in Governance

This section presents the limited evidence available on trends in global averages of governance. Because the means of the governance indicators were rescaled to equal zero in each period, the aggregate indicators are by construction informative about countries' relative positions around the average but uninformative about trends in global averages of governance. To discuss trends in governance worldwide requires going back to the underlying sources of governance data.[22]

For the six dimensions of governance, data are reported from up to four major underlying sources available for each of the four periods 1996, 1998, 2000, and 2002: Economist Intelligence Unit Country Risk Service, DRI Country Risk Review, International Country Risk Guide, and Global Competitiveness Report (table 7). The first three are polls of experts that cover a large set of countries with a consistent methodology from year to year and can therefore be expected to be informative about overall trends. Global Competitiveness Report covers a smaller set of countries, but it is the only survey of individuals that is available in all four periods. To maximize comparability across sources and over time, the focus is for the first three sources on the set of countries common to these three sources for all periods. For the fourth source the focus is on the smaller set of countries available in each period and a small number of survey questions that have been consistently available over time. The underlying data have been rescaled to run from zero to one, and for each source and governance component the score is reported on the same question or average of questions used in the aggregate indicator.

Table 7 reports the average across all countries of each source in each year, the standard deviation across countries for each source, and the *t*-statistic associated with a test of the null hypothesis that the world average score is the same in the first and last periods. Several sources report substantial declines in world averages of the six dimensions of governance. The DRI Country Risk

22. By construction, the standard deviation of the aggregate governance estimates is equal to one in each period, and so these aggregate indicators also cannot be used to assess whether there has been global convergence in governance.

TABLE 7. Global Trends in Governance, Selected Sources

Indicator and Source	Number of Countries	World Average				SD across Countries				*t*-Statistic for Mean Difference from Last Year to First Year[a]
		1996	1998	2000	2002	1996	1998	2000	2002	
Voice and accountability										
DRI	112	—	—	—	—	—	—	—	—	—
EIU	112	0.41	0.42	0.42	0.46	0.30	0.32	0.31	0.28	1.3
PRS	112	0.67	0.66	0.66	0.66	0.23	0.25	0.25	0.25	−0.3
GCS	—	—	—	—	—	—	—	—	—	—
Political stability										
DRI	102	0.82	0.81	0.74	0.70	0.18	0.18	0.24	0.27	−4.1
EIU	102	0.55	0.53	0.58	0.55	0.29	0.29	0.30	0.28	0.1
PRS	102	0.80	0.75	0.74	0.76	0.14	0.19	0.17	0.13	−2.0
GCS	—	—	—	—	—	—	—	—	—	—
Government effectiveness										
DRI	102	0.59	0.58	0.50	0.48	0.27	0.26	0.30	0.30	−2.9
EIU	102	0.41	0.47	0.46	0.41	0.30	0.24	0.24	0.30	0.0
PRS	102	0.63	0.67	0.59	0.67	0.24	0.12	0.10	0.15	1.5
GCS	51	0.45	0.48	0.42	0.31	0.14	0.15	0.11	0.13	−5.2

Regulatory quality									
DRI	106	0.83	0.84	0.79	0.76	0.14	0.14	0.18	0.20	−3.0
EIU	106	—	—	—	—	—	—	—	—	—
PRS	106	—	0.63	0.60	0.76	—	0.20	0.22	0.21	4.6
GCS	51	0.51	0.58	0.59	0.58	0.14	0.15	0.15	0.15	2.5
Rule of law									
DRI	102	0.73	0.73	0.67	0.65	0.20	0.20	0.23	0.24	−2.6
EIU	102	0.50	0.53	0.51	0.54	0.27	0.29	0.30	0.27	1.1
PRS	102	0.77	0.67	0.68	0.66	0.20	0.24	0.22	0.23	−3.8
GCS	51		0.67	0.64	0.59		0.21	0.23	0.24	−1.7
Control of corruption									
DRI	102	0.61	0.60	0.54	0.53	0.25	0.26	0.30	0.31	−1.8
EIU	102	0.37	0.37	0.36	0.38	0.31	0.33	0.32	0.32	0.2
PRS	102	0.62	0.52	0.48	0.42	0.20	0.22	0.22	0.20	−7.2
GCS	51	0.56	0.57	0.59	0.57	0.13	0.14	0.11	0.14	0.3

— Not available.

Note: See table 1 for source codes. Reports trends in cross-country averages of selected components of governance indicators. For EIU, DRI, and PRS, the sample of countries is restricted to those that appear in all three sources in all four periods, to ensure comparability over time and across indicators. For GCS averages are reported across countries of selected individual questions available in all four rounds, with attention restricted to countries available in all four periods.

[a] *t*-Statistic associated with a simple test for equality of global averages in the first and last available periods.

Source: Authors' calculations based on sources listed in table 1.

Review in particular shows statistically significant declines in all five indicators in which it appears. International Country Risk Guide reports significant declines in world averages for political stability, rule of law, and control of corruption but improvements in regulatory quality and government effectiveness. Among polls, the Economist Intelligence Unit Country Risk Service alone consistently does not report any significant trend. Finally, the single survey, Global Competitiveness Report, reports significant deterioration in political stability and government effectiveness and smaller declines in rule of law.

It is not clear how much importance to ascribe to these trends in world averages. On one hand, these statistics represent the only information available on trends over time and so should be taken seriously. On the other hand, the disagreement among sources on the direction of global trends is striking—overall 8 averages improve or remain the same and 11 decline. Looking only at statistically significant changes, however, shows that declines in governance averages outnumber increases 10 to 2 (and both of the statistically significant increases are in regulatory quality).

All that can be cautiously concluded, therefore, is that there is certainly no evidence of any significant improvement in governance worldwide and, if anything, that the evidence is suggestive of a deterioration in key dimensions such as control of corruption, rule of law, political stability, and government effectiveness. It can therefore be safely concluded that the (relative) governance estimates for a country do not underestimate absolute trends because there is no evidence of a worldwide improvement.

VI. Comparisons with Transparency International's Corruption Perceptions Index

Transparency International's pioneering Corruption Perceptions Index (CPI), like the indicators presented here, is an average of ratings reported by a number of perceptions-based sources. In content, the primary differences are that the CPI relies on a subset of the sources used here and it treats multiple years of data from the same source as separate sources in the aggregation procedure.[23] In particular, the 2002 CPI is based on 10 distinct data sources but uses between two and three years of data from some of them and treats them as separate sources, to arrive at a total of 15 components. In contrast, the control of corruption indicator constructed here is based on 14 distinct sources, using only data from 2002 and without using multiple years from the same source.

The CPI also differs in its approach to aggregation (see Lambsdorff 2002 for details). It uses a percentile-matching approach to put data in common units, a simple average of rescaled scores as the estimate of corruption for each country, and a nonparametric bootstrapping procedure (discussed later) to generate measures of precision for the aggregate indicator. The control of corruption indicator, by

23. See Kaufmann and others (2003) for more detailed discussion of the two indicators.

contrast, uses an unobserved-components model to transform individual sources into common units (this is the role of the α's and β's in equation 1), construct an appropriately weighted average of sources as the aggregate score, and produce margins of error to summarize the precision of the estimates of governance.

The estimates of corruption based on the two indicators come to very similar results, with a correlation above 0.9. However, the 2002 CPI covers only 102 countries, because it discards countries with fewer than three data sources. In contrast the control of corruption indicator covers 195 countries, or nearly twice as many. The margins of error generated by the two approaches are similar as well: The average width of a 90 percent confidence interval is 0.94, or 9.4 percent of the range of units from 0 to 10, for the CPI, and 0.71, or 14 percent of the range from -2.5 to 2.5, for the control of corruption indicator. These figures are not fully comparable, however, because the control of corruption indicator covers many more countries, many with only one or two sources of data, and hence would be expected to have somewhat larger margins of error. If only the set of 102 countries appearing in the CPI are included, the average width of a 90 percent confidence interval for the control of corruption indicator is 0.52, or 10 percent of the range of this index—almost identical to the margins of error in the CPI relative to its scale of units.

However, the apparent similarity in the precision of the two indicators is likely to be the result of two offsetting biases in the Transparency International methodology. First, the bootstrapping approach understates the margins of error by overstating the precision of estimates of corruption for countries with relatively few sources. The intuition for this is straightforward. When the number of observations is small, bootstrapped standard errors will understate true standard errors because the observed data are less likely to span the full range of variation in the underlying data generating process. This suggests that the CPI margins of error should spuriously be smaller than those reported for the control of corruption indicator here. In particular, for a country with only three data sources, numerical simulations suggest that the CPI approach will understate the standard error (overstate the precision) of the corruption estimate by about 40 percent.

Second, the estimates of corruption produced by the unobserved-components model used here are a precision-weighted average of individual sources, whereas the CPI approach is based on a simple average. Because precision weighting improves the accuracy of the estimates of corruption, CPI margins of error should correctly be expected to be larger than those reported here. However, according to the estimates in this article, the differences in estimated precision across sources in the CPI are sufficiently small that the benefits of precision weighting are relatively small.

VII. Conclusion

This article presents substantially expanded and updated indicators of six dimensions of governance for 1996, 1998, 2000, and 2002. Large numbers of individual

sources were aggregated, both expanding country coverage and improving the precision of the aggregate indicators. Nevertheless, margins of error remain substantial relative to the units in which governance is measured—especially important to consider when classifying countries according to levels of governance.

An important methodological observation is that there are few alternatives to subjective, experiential data for measuring certain dimensions of governance. Objective indicators of governance, although also very useful, have implicit margins of error on the same order of magnitude as those associated with subjective aggregates. The importance of ideological biases in the perceptions data from polls of experts was empirically investigated and for the most part discounted. Finally, the limited evidence on trends over time in governance worldwide is difficult to interpret, but it can be said with some confidence that there is little evidence of improvements in global governance over the period considered.

As this research project on measuring cross-country differences in governance continues, additional data should become available to enable further improvements in precision. The broader objective is to provide a set of monitorable indicators of governance for individual countries to benchmark themselves against other countries and over time. Limitations will remain, however, in what can be achieved with this kind of cross-country, highly aggregated data. Such data cannot substitute for in-depth, country-specific governance diagnostics as a basis for policy advice to improve governance in a particular country.

REFERENCES

Beck, Thorsten, George Clarke, Alberto Groff, Phillip Keefer, and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1):165–76.

Cingranelli, David, and David Richards. 2001. *Coding Government Respect for Human Rights*. Manual version 1. Binghamton, N.Y.: State University of New York.

Clague, Christopher, Philip Keefer, Stephen Knack, and Mancur Olson. 1999. "Contract-Intensive Money: Contract Enforcement, Property Rights, and Economic Performance." *Journal of Economic Growth* 4(2):185–211.

Di Tella, Rafael, and Ernesto Shargrodsky. 2003. "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires." *Journal of Law and Economics* 46(1):269–92.

Djankov, Simeon, Rafael La Porta, Florencio López-de-Silanes, and Andrei Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117(1):1–37.

———. 2003. "Courts." *Quarterly Journal of Economics* 118(2):453–517.

Efron, Bradley, and Carl Morris. 1971. "Limiting the Risk of Bayes and Empirical Bayes Estimators—Part 1: The Bayes Case." *Journal of the American Statistical Association* 66:807–15.

———. 1972. "Limiting the Risk of Bayes and Empirical Bayes Estimators—Part 2: The Empirical Bayes Case." *Journal of the American Statistical Association* 67:130–39.

Goldberger, Arthur. 1972. "Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables." *International Economic Review* 13(1):1–15.

Golden, Miriam, and Lucio Picci. 2003. "Proposal for a New Measure of Corruption, and Tests Using Italian Data." Working paper, University of Bologna, Italy.

Hellman, Joel, and Daniel Kaufmann. 2004. "The Inequality of Influence." In *Trust in Transition*, ed. J. Kornai and S. Rose-Ackerman. New York: Palgrave-Macmillan.

Kaufmann, Daniel, and Aart Kraay. 2002. "Growth without Governance." *Economia* 3(1):169–230.

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobatón. 1999a. "Aggregating Governance Indicators." Policy Research Working Paper 2195. World Bank, Washington, D.C. Available online at www.worldbank.org/wbi/governance/pubs/aggindicators.html.

———. 1999b. "Governance Matters." Policy Research Working Paper 2196. World Bank, Washington, D.C. Available online at www.worldbank.org/wbi/governance/pubs/govmatters.html.

———. 2002. "Governance Matters II—Updated Indicators for 2000/01." Policy Research Working Paper 2772. World Bank, Washington, D.C. Available online at www.worldbank.org/wbi/governance/pubs/govmatters3.html.

———. 2003. "Governance Matters III: Governance Indicators for 1996–2002." Policy Research Working Paper 3106. World Bank, Washington, D.C.

Knack, Steve, and Mark Kugler. 2002. "Constructing an Index of Objective Indicators of Good Governance." Working paper, World Bank, Washington, D.C.

Lambsdorff, Johann. 2002. "Background Document to 2002 Corruption Perceptions Index." Available online at www.gwdg.de//~uwvw/downloads/2002_cpi_fd.pdf.

Radelet, Steven. 2003. *Challenging Foreign Aid: A Policymaker's Guide to the Millennium Challenge Account*. Washington, D.C.: Center for Global Development.

U.S. Department of State. 2002. "Fact Sheet: The Millennium Challenge Account." Washington, D.C.

Wirtz, Ronald. 2002. "Gross Domestic Product: Understanding News from Noise." *Region*. Federal Reserve Bank of Minneapolis. Available online at http://minneapolisfed.org/pubs/region/02–06/gdp.cfm.