Jnl Publ. Pol., **27**, 3, 255–291
doi:10.1017/S0143814X07000748

Measuring Corruption: A Critique of Indicators in Eastern Europe and Central Asia

STEPHEN KNACK, *Economist, World Bank**

ABSTRACT

Concern about corruption has stimulated the creation of a multiplicity of indicators by a multiplicity of methods by the World Bank, World Economic Forum, Transparency International and commercial rating agencies. However, the construction of indices varies, raising substantive and methodological issues, which are reviewed with particular reference to post-Communist transition societies in Eastern Europe and Central Asia. This paper systematically examines indicators, starting with the definitional distinction between measures of administrative corruption as distinct from state capture. It pays particular attention to problems arising from the aggregation of indicators from multiple sources, and the extent to which the changing composition of cross-country indicators limits trend analysis. It concludes that many of the problems of aggregation may be avoided by using single source and single dimension indexes.

1. Properties of corruption indicators

Corruption is of major importance for governance (Arndt and Oman 2006). There are numerous definitions of corruption in the academic literature and among donor agencies; most are quite broad, and in some cases vague. Transparency International's definition, 'the misuse of entrusted power for private gain', is representative. Often, the term 'misuse' or 'abuse' is further defined to apply only to illegal actions. Given the conceptual definition offered by TI, corruption can be disaggregated along many dimensions. First, one can distinguish between

[1] This paper benefited from valuable suggestions and comments from James Anderson, Nicholas Burger, Cheryl Gray, and Richard Rose, but all remaining errors are the sole responsibility of the author. The conclusions of this paper are not intended to represent the views of the World Bank, its Executive Directors, or the countries they represent.

central, provincial and municipal levels of a political system. Corruption can be widespread at the local government level, even if it is controlled effectively at the central government level. The United States and India provide examples where corruption is much more severe in some states than in others.

Second, a distinction can be made between different broad purposes of the improper actions. For example, bribes may be intended to influence the content of laws and rules, i.e. *state capture*, or alternatively to influence their implementation, i.e. *administrative corruption* (World Bank 2000).

Third, one can distinguish among the actors involved in a corrupt transaction. For example, bribes or diversion of public funds may involve various combinations of firms, households, and public officials. These actors can be distinguished further by their characteristics, e.g. large vs. small firms, rich vs. poor households, low-level vs. high-level officials, roughly corresponding to 'petty' or 'grand' corruption.

Fourth, corruption can be disaggregated by the administrative agency or service involved, such as tax and customs, business licenses, inspections, utility connections, courts, or public education and health facilities. Surveys of firms and households on corruption often emphasize this distinction.

Regardless of one's preferred conceptual definition, the choice of measurement techniques from a limited set of feasible alternatives inevitably produces an implicit definition that can differ substantially from one's ideal. Any pair of assessment methodologies will measure a different (if unknown) mix of these various dimensions of corruption. For example, what weight should be given to central, state and local governments when assessing corruption in federal countries such as the United States or India? What weight should be given to administrative corruption, as opposed to state capture, diversion of funds, or commitment by top officials to fighting corruption?

Table 1 provides examples of different methods for generating country-level corruption measures. The strength of nationally-representative surveys of firms or households is in measuring the incidence of corrupt behaviors encountered by users of government services. This approach emphasizes administrative corruption. However, firm surveys can measure some aspects of state capture, by including questions about improper influence over laws and regulations affecting business. Surveying firms and households is less effective in assessing the prevalence of corrupt transactions that occur entirely within the state, for example when politicians bribe bureaucrats or when funds are illegally diverted. Many types of conflicts of interest also are not easily captured by firm surveys, for example equity stakes of public officials, or employment promises to them by firms (World Bank 2000).

The Business Environment and Enterprise Performance Survey (BEEPS) is a nationally-representative survey of business firms assessing corruption and other problems faced by businesses in the ECA region. The BEEPS is sponsored by the European Bank for Reconstruction and Development (EBRD) and the World Bank, and has covered almost every country in the region in each of three survey waves: 1999, 2002 and 2005. Similar enterprise surveys have been conducted by the World Bank in many countries in other regions, but so far they have been done only on a country-by-country basis rather than region-wide as with BEEPS. Managers of business firms may be viewed as merely a special category of 'well-informed persons'. The distinction nevertheless is important. Questions in the enterprise surveys place a greater emphasis on experience, and less on perceptions.

The World Economic Forum's (WEF) 'Executive Opinion Survey' is another cross-country survey of firm managers. In the 2005 survey, a total of 10,993 responses were received, ranging from 22 for Mauritius to 473 for Russia. Cross-country rankings on several corruption questions (see Appendix 5) are published for 117 countries in WEF's annual Global

Table 1: Major sources of cross-country corruption data

Data sources	Examples
Representative surveys of service users	
Firms	World Bank investment climate assessments (including BEEPS)
	WEF's Executive Opinion Survey
	IMD's executive opinion survey
Households	International Crime Victim Surveys
	New Democracy Barometer, Afrobarometer, Asia Barometer, Latinobarometer
	World Values Surveys
	Global Corruption Barometer (TI)
	Gallup International 'Voice of the People'
Expert assessments	
Experts rating multiple countries	Nations in Transit (Freedom House)
	International Country Risk Guide (ICRG)
	Economist Intelligence Unit (EIU)
	Global Insight
	World Bank CPIA
Surveys of 'well-informed persons' within country	UNECA African Governance Indicators
	World Governance Assessments
Composite indexes	
Aggregation from various sources	TI perceptions of corruption index
	WBI control of corruption index

Competitiveness Report (Lopez-Claros, Porter and Schwab 2005). Ratings are computed as the simple average of all executives' responses. The Swiss IMD (Institute for Management Development), uses a nearly identical methodology, but somewhat different survey questions in its World Competitiveness Yearbook (IMD 2005). The IMD executive survey is conducted in many fewer countries than the WEF survey, and includes fewer questions on corruption. The IMD also discloses less information than the WEF on the size and composition of its sample of executives in each country.

The WEF and IMD executive opinion surveys differ from the BEEPS (and the World Bank's other firm surveys) in several important respects. First, the sample in each country is selected with a preference for executives with international experience, who tend to be from larger and exporting firms. Second, the questions are designed to elicit 'the expert opinions of business leaders' on corruption and other issues, and focus much less than BEEPS on firms' experiences. The WEF, for example, asks about diversion of public funds, an issue on which few firms would have direct knowledge. Third, the WEF and IMD surveys are designed solely to produce country-level measures of the business climate. The BEEPS (and other World Bank firm surveys) is designed for firm-level analyses, and the datasets include numerous characteristics of the responding firms, while taking care to preserve firm anonymity to encourage candid responses.[1]

Household surveys addressing corruption issues are not quite so well developed as firm surveys. Beginning in 2003, Transparency International has sponsored the annual Global Corruption Barometer (GCB), conducted with assistance from Gallup International's survey network. The World Values Surveys (WVS), International Crime Victimization Surveys (ICVS), 'Voice of the People' surveys by Gallup International, and several regional 'Barometer' surveys have also included questions on households' experiences with or attitudes toward corruption.

Expert assessments of corruption provided by various commercial and other organizations have been most widely used for comparisons across countries and over time. Their methods differ in several potentially important ways. First, they differ in the degree to which assessments are 'centralized'. The centralized type is exemplified by Nations in Transit (NIT) and by the International Country Risk Guide (ICRG). Corruption ratings from these sources are formed by a network of correspondents with country-specific expertise, but the final ratings are determined centrally by a very small number of people. In the decentralized type, views are solicited from experts only for countries in which they have direct experience. The Africa Governance Indicators (AGI), of UNECA covering corruption and other governance issues, are based on surveys of

elites in 28 countries, conducted in 2002–2003 (Economic Commission for Africa 2005). The AGI 'expert panels' varied in size from about 70 to 120 across countries.[2] World Governance Assessments were conducted in late 2000 and early 2001 in 22 developing countries from various regions, including Bulgaria, Kyrgyzstan and Russia (Hyden, Court and Mease 2004). In each country, 35 'well-informed persons' were asked 30 questions, including three pertaining to corruption in business licensing, in the judiciary, and favoritism in applying regulations. Data in six of the 22 countries were deemed to be of unacceptably low quality, so the publicly available data set covers 16 countries.[3]

The World Bank's Country Policy and Institutional Assessment (CPIA) is a hybrid of centralized and decentralized expert-based ratings. The ratings originate with the country teams and regional offices, but then are reviewed for cross-regional comparability by central units. Most ratings proposed by the regions are not changed in this review, however, and the final ratings are correlated at about 0.98 with those proposed by the regions.

Secondly expert assessments differ in the extent of documentation they provide regarding definitions and methods. For example, Nations in Transition (Appendix 2) provides more details than ICRG on its assessment criteria and its methodology (including sources of information), and provides extensive country narratives containing qualitative assessments of corruption problems to accompany the quantitative ratings. The CPIA is transparent in some respects but opaque in others. Its detailed assessment criteria are posted on a public web site, and there are reasonably detailed narratives justifying the ratings. However, these justifications are not publicly released, and the ratings themselves are available only for the IDA-eligible (mostly low-income) countries. Sources that are more transparent and accountable, as reflected by the availability of detailed assessment criteria and justifications for ratings of each country, arguably will tend to be more accurate in their assessments. It is impossible to debate meaningfully the appropriateness of country ratings when definitions are brief, vague, and broad, and when ratings are not accompanied by justifications for each country.

Corruption indicators also differ in attempting to assess (a) the relative incidence of corrupt transactions, (b) the impact of corruption on business, or (c) the existence of laws, policies and associated enforcement mechanisms believed to affect the prevalence of those transactions. The ICRG is an example of type (a), while type (b) is illustrated by the NIT corruption index (Appendixes 2 and 4). The World Bank's CPIA question 16 (see Appendix 3) is a mix of types (a) and (c). Most questions in the BEEPS and WEF are of type (a), but each source contains type (b) questions also. One BEEPS question asks how problematic is corruption

'for the operation and growth of your business'. Two other BEEPS questions ask about the 'impact on your business' from other firms' payments to parliamentarians or government officials to influence laws and regulations. The WEF similarly asks whether or not 'other firms' illegal payments to influence government policies, laws, or regulations impose costs or otherwise negatively affect your firm'.

Sources producing corruption indicators have different constituencies or audiences, with potential implications for what their ratings are measuring. Some sources, such as Freedom House, which produces Nations in Transit, are advocacy NGOs. Others, such as the ICRG, are marketed by profit-making companies to multi-national investors and other paying subscribers. Most subscribers to the ICRG are more interested in conditions facing foreign investors than in those facing local investors. To the extent corruption-related obstacles differ for those two sets of investors, the ICRG ratings can be expected to focus on those most pertinent to its paying subscribers. Corruption ratings produced by development agencies, including the World Bank's CPIA, and similar ratings produced by the African Development Bank and Asian Development Bank, are also potentially influenced by their constituents. Because the CPIA ratings are important in determining aid allocations for the World Bank's lower-income (IDA-eligible) countries, the Bank's country teams could benefit from proposing higher-than-warranted ratings. Providing more favorable assessments can also make working relations of country teams with their government counterparts easier. However, statistical analysis finds no evidence that IDA-eligible countries are overrated relative to non-IDA countries. Specifically, if the CPIA corruption ratings are regressed on other available corruption indicators and on a dummy for IDA eligibility, the coefficient for the latter is negative, instead of positive as implied by the potential incentive bias.

Corruption indicators differ in conceptual breadth; some have more dimensions than others. The ICRG, NIT, and CPIA each provide a single measure of corruption intended to reflect a mix of various aspects of corruption. The BEEPS and WEF surveys contain multiple questions pertaining to narrower aspects of corruption. For some purposes, broader measures may be preferred: a researcher testing the hypothesis that more women in parliament reduces corruption (Swamy et al. 2000), or that corruption slows economic growth (Mauro 1995) may not be concerned about exactly how corruption is defined. Theory may provide little guidance as to which aspects of corruption are most harmful to growth. Similarly, a donor wanting to direct more aid to less corrupt countries may have no particular view on which aspects of corruption most impair aid effectiveness. For other purposes, however, narrower measures may

be required. For example, an effective and convincing test of the hypothesis that higher civil service pay reduces bribe-seeking may require measures of administrative rather than grand corruption. A donor funding projects in a country may be interested in a measure of corruption in public procurement, while a donor providing budget support might prefer a measure of the likelihood of diversion of funds to unintended purposes. The design of effective anti-corruption reforms requires narrow measures to identify specific problem areas and track progress over time.

Broader corruption measures such as the ICRG or NIT not only are less conceptually precise for good or ill, but their meanings also tend to be more *uncertain*. For the ICRG, NIT or CPIA corruption indicators, the weights given to the various aspects of corruption listed in their assessment criteria are unknown, by contrast with the multi-item index from several BEEPS or WEF corruption measures. Aggregation implies a reduction in conceptual precision, but there is no increase in uncertainty over what is being measured if the data user selects the indicators to include in the index and the weights assigned to each indicator. However, with broader, multi-dimensional indicators such as ICRG, data users have no way of knowing exactly what the indicators are even attempting to measure. This uncertainty problem is exacerbated when no such criteria are made public, as is the case for corruption measures produced by two competitors of the ICRG, the Economist Intelligence Unit (EIU) and Global Insight.

Finally, some corruption indicators differ in their ability for measuring changes over time. Broad, multi-dimensional indicators are potentially problematic in this respect, because there is no way to ensure that the implicit weights given to the various dimensions do not vary over time. Some indicators do not have fixed and explicit criteria provided for each ratings level, so there is no way of ensuring that a rating of, say, 4 means the same thing from one year to the next. The ICRG is an illustrative example. Its ratings guide (PRS Group 2003) states that ratings are intended to be comparable both across countries and over time. But it provides no indication of what conditions are described by a rating of 2, 3, 4, etc. Nations in Transit provides only a generally worded set of criteria for each of its 1 to 7 ratings levels, written to apply not only to corruption but to NIT's six other indicators.[4] The WEF questions on frequency of irregular payments have 7 response categories, ranging from 'is common' (1) to 'never occurs'. How respondents interpret 'common' may be relative. In principle, the CPIA criteria are fixed and explicit, but in practice they are revised somewhat every few years, and they are sufficiently subjective that the standards for a given ratings value may not be fixed.

Changes in methods, as well as in content, can reduce over time comparability of indicators. Admirably, WEF has tried to increase the response rate of its Executive Opinion Survey, to enhance accuracy by making the sample more representative. The mean number of responses per country increased from 84 in 2004 to 94 in 2005. However, progress on this front can affect apparent trends. Suppose executives with the strongest opinions are the most likely to respond, and that strong opinions tend to be unfavorable. An increase in the response rate from one year to the next would then reduce the negative bias, but the year-on-year *change* would be biased toward showing improvement.

2. *Composite corruption indexes*

The first motive for constructing a single corruption index from multiple, distinct sources of corruption indicators emphasizes substantive content: individual indicators, or even several indicators from one source such as the BEEPS, may be defined too narrowly for certain purposes. For example, no matter how many corruption indicators one aggregates from the BEEPS, the resulting index still reflects only corrupt interactions between firms and public officials.

The second motive is to reduce measurement error. Given the obvious difficulties in measuring corruption, any one source may be highly inaccurate. However, *if* errors in measurement are largely independent across sources, the errors will tend to cancel out when data are aggregated from multiple sources. The third motive is to cover a larger number of countries. No one source covers all countries. Some sources do not overlap at all in country coverage, for example the UNECA's African Governance Indicators and Nations in Transit.

The latter two motives were responsible for the creation of Transparency International's widely-cited 'Corruption Perceptions Index', and subsequently WBI's 'Control of Corruption' index (Kaufmann, Kraay and Mastruzzi 2006). Although the statistical methods vary somewhat, both standardize corruption indicators from numerous sources to place them on a comparable scale, and compute an average (unweighted for TI, weighted for WBI) to obtain one value for each country. Missing values on any indicator for a given country are ignored, so are in effect imputed as the average of all indicator values for which data are available for the country. By this procedure, an index value can be computed for any country for which data is available from even one of the many sources used.

The original purposes of the TI index were to raise awareness of corruption, and to provide researchers with better data for analyzing the causes and consequences of corruption. The WBI index, appearing several years after the TI index, was intended by its authors to expand on

TI. First, the WBI index provides a value for any country with data available from even one source, while the official TI index requires three sources. Second, the WBI index incorporates data from more sources, including ICRG and others which TI rejects (Lambsdorff 2005a). Third, using many of the same data sources, WBI constructs five other broad 'governance' indexes, titled Rule of Law, Voice and Accountability, Political Stability and Violence, Regulatory Quality, and Government Effectiveness. Fourth, WBI weights available sources differently, in contrast to the equal weighting in TI of available sources for each country. Finally, WBI attempts to improve on the treatment of statistical uncertainty in TI. While TI lists number of sources, and the range and standard deviation among sources, WBI computes a 'standard error' as an indicator of uncertainty accompanying each point estimate. These standard errors are lower for countries covered by more data sources and for countries covered by data sources which are more highly correlated with other sources in the index.

For the consciousness-raising and research purposes that inspired these aggregate indexes, the intuition underlying them is plausible. Measurement error is likely to be reduced somewhat by combining data from multiple sources. The expansive definition of corruption implied by aggregation was a virtue for TI's and (later) the World Bank's consciousness-raising agendas, and for cross-country empirical research demonstrating adverse economic consequences of corruption. The limitations of these composite indexes are often neglected by data users, however (Arndt and Oman 2006). Some of these problems are common to the broad corruption measures from individual sources such as ICRG, NIT or CPIA. Other limitations are introduced by the process of aggregation.

Transparency in construction

If any component of a composite index is constructed in an opaque manner, the composite index in turn will be somewhat opaque, regardless of the transparency of the aggregation process itself. If the documentation in the ICRG, for example, provides little guidance as to how various aspects of corruption are weighted, or what information sources are used, one cannot fully explain what the WBI 'Control of Corruption' index is measuring or on what basis. Although both TI and WBI provide thorough explanations of their aggregation methodology, replication of the indexes by independent analysts would be costly, particularly as the number of sources used has expanded over the years. Some of the sources are available only to paying subscribers or members, and some are not publicly available at all.

Conceptual imprecision, uncertainty and inconsistency

The TI and WBI indexes are conceptually more *imprecise* than some of their broadly-defined components (e.g. ICRG, NIT and CPIA) also more conceptually *uncertain* in how criteria are weighted. In contrast to any single broadly defined indicator, the TI and WBI composite indexes suffer from having *varying* definitions. Composite indexes have no explicit definition; instead, they are defined implicitly by what goes into them. The sources used in constructing these composite indexes change over time, so the implicit definition of corruption reflected in the index changes over time. Moreover, the sources used in constructing the indexes vary from country to country in a given year.

Among the 27 ECA countries, there are 13 distinct combinations of sources used in computing the 2005 TI index, so the 27 index values reflect 13 different implicit definitions of corruption. Index values for the three Baltic countries are based on three distinct combinations of indicators. Values for Bulgaria, Romania, and Croatia – which like the Baltic nations are often compared to each other – are also based on three different combinations of indicators. The same is true for the three Caucasus countries, Armenia, Azerbaijan and Georgia. This comparability problem is even more severe for the 2004 WBI index on Control of Corruption. It uses 23 different combinations of sources for the 27 ECA countries. No one combination of sources is used to construct index values for even three countries. There are only four pairs of countries whose values are based on a common set of sources: Russia and Poland are based on the same 14 sources, Estonia and Romania on the same 13 sources, Bulgaria and Lithuania on the same 12, and Croatia and Latvia on the same 11.[5]

In principle, more strictly comparable comparisons could be performed simply by computing a composite index that deletes any source not common to the two countries in question. Alternatively, one could compare two countries source by source, not bothering to construct a composite index at all. Either of these options requires access to the complete underlying data, however, which neither TI nor WBI provide in full.

Tracking changes over time

The standardization procedure used to place different indicators on a common scale precludes the ability to track changes in ranking meaningfully over time (Arndt and Oman 2006). The WBI index, for example, is constructed to have a mean of 0 and a standard deviation of 1 for each year the index is provided. Neither index values nor rankings are

comparable across years when the composition of the sample changes. The addition of Luxembourg to the TI sample in 1997, and Iceland in 1998, reduced the rankings of most other nations. This limitation of the composite indexes often is not appreciated, as reflected not only in numerous media references to the TI index but also in many internal World Bank memos, and even in papers submitted for publication to academic journals.

The over-time comparability problem raised by changes in country coverage can be corrected, for the most part, by comparing rankings over time for a constant set of countries. For example, among the 102 countries included in the TI index in every year between 2002 and 2005, Slovenia's rankings were 27th, 26th, 28th and 27th. Neglecting to adjust for a common sample, its ranking falls from 27th to 31st.

The above method corrects only for changes in coverage but not for year-to-year changes in the underlying data sources and indicators available for the country in question. For example, no TI index value for any ECA country was based on the same set of sources in both 2004 and 2005. The WBI indexes for 2002 and 2004 are based on the same set of sources for only 4 of the 27 countries in the ECA region: Hungary, Slovak Republic, Slovenia, and Tajikistan. As with a pair-wise comparison of countries at a point in time, a comparison for a single country at two points in time would be more convincing if it were based on a common set of sources. Again, one could do this in principle, by going to the component data sources, but many would be costly or impossible to access for most data users. A second-best solution would be for TI and WBI to add to their web sites a tool that allows purer comparisons over two time periods (or across two countries) by computing customized indexes based only on sources common to both years (or countries).

A potentially more serious problem is that changes in perceptions of corruption may lag reality, if they have anything to do with reality at all. Peru's ICRG ratings have been lower since Fujimori and Montesinos left office than they were during their rule. Similarly, Ireland's ICRG rating was a very favorable 5 throughout the years Charles Haughey was Prime Minister, but declined to a low 2 in 1997 – years after his retirement from politics – with revelations of his corrupt activities while in office.

Recognizing some of these methodological issues, Lambsdorff (2005a, 2005b) is careful to interpret changes in the TI index as shifts in *perceptions* of governance. Kaufmann (2005) on the other hand interprets statistically significant increases in the WBI indexes as improvements in *governance*, demonstrating that 'countries can substantially improve' their quality of governance 'even in the short term'.

Interdependence of sources

Intuitively, if several sources assess a country more favorably in year 2 than in year 1, we can infer more confidently that an actual improvement occurred than if evidence of progress were based on a single data source. This intuition is valid only to the extent that different sources represent independent judgments. In classifying which countries have improved or worsened to a 'statistically significant' degree over time, both WBI and TI assume that assessments from each source are fully independent. However, many of their sources clearly are not independent (Arndt and Oman 2006). The CPIA process takes into account numerous expert assessments and firm surveys, and ratings often are adjusted to be more consistent with rankings from those sources. The expert assessments of the 'centralized' type in turn often consult each other, and sometimes adjust ratings for outliers. The EIU provides little information on definition or methodology for its corruption rating, as noted by Lambsdorff (2005b). He shows that the EIU ratings are strongly related to lagged, but not contemporaneous WEF corruption ratings. The simplest explanation for this result – although not one mentioned by Lambsdorff – is that the EIU assessments may systematically incorporate the most recently available WEF ratings.

Both the empirical evidence and economic incentives suggest that governance ratings such as the CPIA and ICRG will take into account any available inexpensive sources of information, including the widely-disseminated TI rankings. The organizations producing these ratings have no interest in throwing away public information from other sources, merely because doing so would benefit others who produce aggregate indexes with explicit weights such as TI or WBI. It would be irresponsible to their paying subscribers if ICRG, EIU, and Global Insight ignored each others' assessments and those of readily available aggregates such as the TI index. Probably for this reason, representatives of these three organizations all readily acknowledge consulting the TI index and/or their competitors' ratings in formulating their own.

Interdependence does not even require, however, that sources directly check each other ratings. It can also result merely from sources relying on many of the same media reports or other qualitative sources of information about conditions in countries. If, for example, the international media cover corruption incidents more thoroughly for large and more newsworthy countries than for small countries, the same unintentional bias against large countries could show up in corruption ratings by different sources.

In contrast to most expert assessments, surveys of firms and households generate data likely to be largely independent from other judgments.

Most respondents in business surveys such as the BEEPS are unlikely to know the TI ratings for the country in which they operate, and even for the few that do know, it is unlikely to influence their response to a question on the share of their firms' revenues paid in bribes. The WEF 'Executive Opinion Survey' differs from BEEPS in several respects that could make it less independent. First, the sample of executives is deliberately chosen to elicit the views of 'business leaders' with extensive international experience. These executives are more likely than those in the BEEPS to be aware of the TI and other cross-country ratings. Second, the WEF survey questions are deliberately phrased in such a way that respondents will 'compare their own environment to a world standard, rather than thinking in national terms'. Some respondents may consult other cross-country rankings in order to provide a seemingly better informed response. Third, the WEF and IMD both implement similar executive surveys, with samples selected by 'partner institutes'. The WEF and IMD share many of the same partner institutes: in 12 of the 51 countries included in both sets of surveys, the WEF and IMD have at least one partner institute in common. Thus, many of the same executives are likely to be included in both sets of surveys.[6]

It is impossible to determine quantitatively the degree of interdependence among sources used in the TI and WBI indexes. Kaufmann, Kraay and Mastruzzi (2006) claim that correlated errors among expert sources are likely to be minimal, because expert ratings are no more correlated with each other than they are with survey sources. This assertion is misleading, however, as it holds only for the WEF firm surveys. The various other household and firm survey indicators tend to show much lower correlations with expert ratings. Moreover, the WEF is an influential component of the TI index, and expert ratings such as ICRG are known to consult the TI index in making their own corruption assessments.

Many of the cross-country or over-time differences they classify as 'statistically significant' would not be, if the appropriate corrections for interdependence could be made. The unknown but substantial degree of interdependence among many of the sources also negates claims regarding the 'precision' of estimates. Other things equal, one can have more confidence in a rating based on 9 sources than on a rating for another country based on only 3 sources. It is also important however to identify the sources and to consider the likely degree of interdependence among them. Three sources consisting of a firm survey, a household survey and an expert assessment can provide a richer set of information than nine sources, if all nine are expert assessments. Iceland's 2002 TI index is computed from six sources, which at first glance appears impressively diverse. However, none of them are truly independent: three of them are

from WEF surveys for 2000, 2001 and 2002, and the other three are from IMD surveys for the same years.[7] Although the partner institutes in Iceland are different for WEF and IMD, the likelihood of overlapping samples of top executives with international experience in a country so tiny must be very high. Iceland in TI is an extreme example of interdependence, but the problem in more moderate form is endemic to both TI and WBI.

The choice of weights in aggregation

Simplicity, objectivity, transparency and replicability all argue for weighting equally each variable or each source, for sources can provide multiple indicators, in constructing a composite index. The TI index weights each of its sources equally, with a caveat: the three most recent WEF and IMD surveys are each included as a separate source. They each therefore receive triple the weight given to another source, such as the EIU or Global Insight.

The goal of accuracy could justify differential weighting, if there is good reason to believe that some sources are more informative than others. The WBI index weights some sources more heavily than others. Sources that tend to be more highly correlated with the other sources are given greater weight, with the precise weights determined objectively by a variant of principal components analysis. The assumption is that if sources are independent of each other, a source that agrees less with the others is a less accurate measure of corruption, whether due to pure measurement error (the source is deficient in measuring what it purports to measure) or due to extraneous content (a source's assessment criteria include factors other than corruption). The rationale for such a procedure disappears however if measurement error is correlated among sources, i.e. if they are not independent. If high correlations among expert assessments are driven by the fact that they consult each other's ratings – or even by experts all basing their ratings on the same information sources – agreement among them is a dubious proxy for accuracy. In that case, any truly independent source, using different information or a different methodology, is likely to generate ratings less correlated with the interdependent expert ratings than the latter are with each other.

The BEEPS is a good illustration of this problem. In the WBI indexes for 2002 and 2004, the weight given to Nations in Transit (covering mostly the same countries) is 24 times the weight given to the BEEPS. As one 'expert'-based source among many in the index, it is unsurprising that Nations in Transit tends to be more highly correlated than a firm survey with most other sources. More defensible than the assumption that

all sources are independent would be an assumption that the *types* of sources listed in Table 1 are largely independent. This more conservative assumption would suggest giving equal weight to each type of source available for a given country, e.g. one-quarter each to firm surveys, household surveys, decentralized and centralized expert ratings.

Interdependence of expert sources can even undermine the main premise of the WBI index methodology that more information – more sources – produces more accurate and reliable estimates. The addition of another expert-based source containing little new information – relying on the same information sources as its competitors, or even checking their ratings – can reduce accuracy of the composite index, by further reducing the weight given to the few sources that do provide truly independent information.

The availability of the composite indexes themselves can aggravate problems. Some expert-based sources providing broad assessments of corruption may, sensibly enough, agree with the premise underlying the TI and WBI indexes that more information is better, and adjust their ratings to conform better to the composites' rankings. The ICRG appears to have done this in late 2001. It publishes monthly corruption ratings, but in most months very few ratings are changed for the 140 countries covered. September to October 2001 was typical, with a single half-point change for Switzerland. From October to November, however, 47 ratings were reduced and 10 increased. Month-to-month correlations in ICRG's corruption ratings always exceed 0.99, but fell to 0.88 for October and November 2001 ratings. The ICRG did not respond at the time to requests for an explanation for this break in the data, but it appears that ratings were re-adjusted to conform much more closely to the TI rankings.[8] The ICRG ratings were correlated with the TI 2001 ratings (released at end of June 2001) at only 0.72. However, the correlation with TI rose to 0.91 with the massive re-calibration by ICRG in November. This evidence of interdependence between TI and ICRG does not directly present a problem for the TI index, which does not include ICRG as a source. It does imply a circularity problem for WBI, however, which uses ICRG and most of the TI sources. It also indirectly suggests a problem for TI and WBI, to the extent that ICRG may not be unique among sources in free riding on the assessments of other sources – including the TI and WBI indexes themselves – rather than basing assessments on their own independent information.

This logic is not unique to governance measurement issues. Surowiecki (2004) makes an analogous argument that more efficient dissemination of information can contribute to stock market bubbles and crashes. A 'herd mentality becomes endemic' as investors cease to make independent judgments about asset values, and the efficiency gains from aggregating

information from large numbers of investors is lost. Similarly, producers of data such as the CPIA and ICRG may devote fewer resources to generating private information on the quality of governance when more public information is made available via widespread dissemination of the TI and Kaufmann-Kraay indexes – even if doing so reduces the accuracy of subsequent iterations of those aggregate indexes.

The interdependence of information on governance undermines the basis for the WBI weighting method. Fortunately, correlation with other sources is not the only proxy for accuracy that could be used in assigning weights in construction of a broadly-defined composite index of corruption. For example, one could weight more heavily those sources that represent truly independent assessments – the BEEPS would thus receive a greater weight than WEF or ICRG. Weighting each type of source equally, as suggested above, is consistent with this reasoning. Or, one could weight more heavily those sources with more extensive publicly available documentation, particularly regarding assessment criteria and methodology and detailed justifications. Nations in Transit would thereby receive a greater weight than EIU. Among the survey sources, one could weight more heavily those with larger and more nationally representative samples, and those that include more questions on corruption. The WEF – with many more corruption questions – would thus be weighted more heavily than the IMD.

A disadvantage to most weighting schemes is that weights would be determined subjectively, in contrast to the objectively-determined weights in the WBI methodology. The larger point is that no one weighting choice is likely to be the most appropriate for all purposes to which an aggregate index might be applied. Greater public access to the underlying data used in the TI and WBI indexes, along with better information on how those underlying data are generated, would permit data users to customize their own indexes more appropriate to their own purposes.

3. Levels and trends in corruption for ECA countries

Table 2 reports summary statistics for the corruption variables included in both the 2002 and 2005 BEEPS. Figures reported represent means, weighting each of the 27 countries equally. The full wording of the survey questions is reported in the Appendix. Note that questions are phrased in terms of bribes typically paid 'by firms like yours', to elicit more candid responses than if respondents were asked directly about bribes their own firm had paid.

The most dramatic improvement between 2002 and 2005 is in the 'bribe tax', which fell by one-third from 1.6 per cent of firm revenues to

1.1 per cent. The bribe tax reported is skewed across firms, with a majority of firms reporting 0 per cent in both years. A positive value for 'bribe tax' was reported by 44 per cent of firms in 2002, declining to 37 per cent in 2005.

Among the numerous other questions on corruption issues in the BEEPS, most show evidence of modest improvement. For example, corruption was cited as a major or moderate obstacle to doing business by 21 per cent of firms on average in 2002, falling to about 18 per cent in 2005. About 26 per cent of firms on average in 2002 reported that paying bribes was frequently, usually or always necessary 'to get things done with regard to customs, taxes, licenses' etc., down to 20 per cent in 2005. Most questions about specific public services also show evidence of declines in the incidence of bribe paying, e.g. in getting connected to public utilities, in obtaining licenses, and in paying taxes and customs.

There is little evidence of change overall in three survey items on 'state capture'. Paying bribes 'to influence the content of new legislation, rules or decrees' appears to be about equally common in both years. Similarly,

Table 2: BEEPS 2002–2005 Summary statistics

	2002	2005
	% of sales paid in bribes (mean)	
Bribe tax (Q40)	1.59	1.05
Kickback for government contract (Q42)	1.99	1.85
	corruption moderate or major obstacle (% of firms)	
Corruption problematic for business (Q54q)	21.0	17.8
	bribes required frequently, usually or always (% of firms)	
Bribe frequency (Q39a)	25.6	25.6
Bribe predictability (Q39b)	30.0	30.0
Utilities (Q41a)	6.3	6.3
Licenses & permits (Q41b)	15.9	15.9
Government contracts (Q41c)	15.9	15.9
Health/safety inspections (Q41d)	10.8	10.8
Fire/building inspections (Q41e)	11.4	11.4
Environmental inspections (Q41f)	7.9	7.9
Taxes & tax collections (Q41g)	17.5	17.5
Customs/imports (Q41h)	15.4	15.4
Courts (Q41i)	8.4	8.4
Influence legislation/rules (Q41j)	5.0	5.0
	moderate, major or decisive impact on firm (% of firms)	
Impact of capture: parliament (Q44a)	10.1	10.1
Impact of capture: govt. officials (Q44b)	12.4	12.4

See Appendix 1 for full wording of BEEPS survey questions. All percentages reported in table weighting each country equally.

the share of firms reporting a significant impact on their business from Parliamentarians receiving bribes to affect their votes is little changed. A slight improvement is evident for a similar question on payments to government officials to affect the content of government decrees.

A few corruption items in the BEEPS show slight deteriorations over time for the region overall. Bribe-paying in obtaining government contracts and in dealing with courts appears to have increased very slightly between 2002 and 2005.

Overall, Table 2 indicates notable progress between 2002 and 2005 in administrative corruption, but not in state capture. Moreover, while most areas of administrative corruption show improvement, progress appears to be uneven and even absent in a couple of important areas, such as the courts.

Overall progress hides uneven progress between countries. Detailed country-by-country results are reported in Anderson and Gray (2006), which also contains several short case studies on corruption successes (Georgia) and failures (Kyrgyz Republic). The multidimensionality of corruption is apparent from these results, suggesting the difficulty in concluding that 'corruption is worse' in country X than in country Y. For example, Macedonia ranked fifth-best in the region on 'bribe tax' in 2005, but worst among all 27 countries on corruption as an obstacle to doing business. Latvia ranks fourth best on bribe frequency, but twentieth on state capture. However, the various measures are significantly and positively correlated, and there are some countries that rank consistently high (Estonia, Slovenia) or low (Albania, Azerbaijan, Kyrgyz Republic) across measures.

As for changes over time, different measures often move in opposite directions within a given country. Georgia is among the few countries showing largest improvements on most corruption questions in the BEEPS between 2002 and 2005. Other striking cases of improvements in corruption are Slovak Republic, Romania and Bulgaria. Slovenia and Estonia also show impressive improvement, given that they already had relatively low levels of corruption in 2002. Azerbaijan and Lithuania exhibit increasing corruption on a range of questions. Kyrgyz Republic's deterioration on many questions must also be considered disappointing, despite a large improvement in the bribe tax from a region-worst 3.7 per cent in 2002 to a second-worst 2.5 per cent in 2005.

Trends in other data sources

Do other sources tend to corroborate the generally favorable trends in corruption shown in the BEEPS? Three key distinctions should be kept in mind in comparing trends from BEEPS to trends in other assessments

of corruption. First, most other sources do not 'unbundle' corruption across various functions of government, but provide only a single broadly-defined indicator. Second, BEEPS measures only corruption in firm-state interactions, while other broadly-defined indicators will also reflect corruption in household-state interactions, diversion of public funds, etc. Third, most other sources are designed primarily to compare corruption levels across countries, and only secondarily to compare corruption levels over time within countries. Although such sources are not very informative on whether corruption is improving or deteriorating for ECA or other regions, they can still be used to compare *relative* performance. Namely, they can help answer the question of whether ECA overall is improving relative to other regions.

Nations in Transit (NIT) covers only the 27 transition countries in ECA; Turkey is excluded. On the 1–7 NIT corruption scale, a 1 is the best possible rating and a 7 is the worst, with quarter-point increments allowed. The mean rating improved from 4.85 in 2002 to 4.80 in 2005. This small average improvement hides substantial variation however: ratings improved for 10 countries, mostly in Eastern Europe, and deteriorated for seven others, mostly in the former Soviet republics, although the two largest, Russia and Ukraine, show small improvements.

The CPIA question, 'Transparency, Accountability and Corruption in the Public Sector', is assessed on a 1–6 scale for 27 ECA countries. The mean value in 2002 was 3.11, increasing to 3.30 in the 2004 ratings.[9] As shown in Table 4, most regions show modest improvement over time, but the increase for ECA was exceeded in magnitude only by the East Asia and Pacific region. Among all 134 countries in the CPIA in both 2002 and 2005, the average ranking for ECA countries was sixty-fourth in 2002, improving to sixty-first in 2005. In 2002, the mean rating for ECA was third-highest among regions, behind Latin America and South Asia. In 2004, ECA ranked behind only Latin America.

The International Country Risk Guide (ICRG) rated 140 countries, including 21 ECA countries, in June 2002 and March 2005. The ICRG is updated monthly, and data for those months were selected to coincide with the beginning of fieldwork for the BEEPS II and III. Unlike the CPIA, the ICRG sample includes most developed countries. Its corruption ratings range from a minimum value of 0 to a maximum of 6. The mean ECA rating increased from about 2.1 in 2002 to 2.2 in 2005. The average ranking for the 21 ECA countries also improved over the period, from 82 to 76.

The World Economic Forum data (WEF) included only 14 ECA countries among a total of 79 with data available for both 2002 and 2005. These 79 include many developed countries. Table 3 reports on nine WEF variables, all scaled from a low value of 1 to a high value of 7.

Trends are highly mixed. The first four in the table, pertaining mostly to state capture, all show either stagnation or deterioration. In particular, the average rating for ECA countries on 'business costs of corruption' – defined in terms of 'other firms' illegal payments to influence government policies, laws or regulations' – worsens from 4.5 to 4.1. The average ranking for ECA on this question fell from 45 in 2002 to 51 in 2005.

Trends are much more favorable on five measures of administrative corruption in the WEF. The average ECA *ranking* improves on all five of these measures, although its average *rating* on the 7-point scale fell from 4.8 to 4.4 on one of them, 'irregular payments in judicial decisions'. This evidence from the WEF is remarkably consistent with the BEEPS in two major respects. First, there is evidence of improvement in administrative corruption, but not in state capture. Second, both sources 'unbundle' administrative corruption in similar ways, finding more evidence of improvement for certain functions (licenses and permits, tax and customs, utilities) than for others (public contracts, judicial system).

Table 3: Corruption trends for ECA in non-BEEPS sources

Source	Sample ECA, all	Indicator scale	Mean value 2002	Mean value 2005	Mean rank 2002	Mean rank 2005
CPIA Q16	27, 134	1–6	3.11	3.30	64	61
Nations in Transit*	27, 27	1–7	4.85	4.80	—	—
ICRG	21, 140	0–6	2.07	2.19	82	76
EIU	20, 115	1–5	2.10	1.95	54	49
WEF	14, 79	1–7				
Favoritism in decisions			2.91	2.79	52	53
Diversion of public funds			3.29	3.33	48	48
Business costs of corruption			4.51	4.07	45	51
Financial honesty of politicians			2.11	2.17	50	50
Irregular payments in . . .						
exports & imports			4.44	4.77	49	45
public utilities			5.06	5.38	46	42
tax collection			4.89	5.29	44	41
public contracts			3.78	3.90	49	48
judicial decisions			4.80	4.42	51	49
TI corruption perceptions	28, 152	0–10	3.25	3.29	92	88
WBI control of corruption	28, 196	0 ± SD	− 0.43	− 0.32	120	110

Second column indicates number of countries, in ECA region, and overall, covered by each data source. Third column shows minimum and maximum possible values of indicators. Columns headed 'mean value' show change from 2002 to 2005 for ECA countries, weighting each country equally. 'Mean rank' shows changes from 2002 to 2005 in mean rank for EAC countries, within the full samples covered.

*For NIT and EIU, larger values indicate more corruption. Appendix provides definitions for the various indicators. Median value for NIT was 5.25 in 2002, improving to 5 in 2005. Average change in CPIA Q16 is for 2002–2004.

The Economist Intelligence Unit (EIU) assigns countries to one of five categories, with a 1 for the least corrupt, and a 5 for the most corrupt. The average for 20 ECA countries included in the 2002 and 2005 ratings improved from 3.10 to 2.95. In contrast, the average for 63 other World Bank borrowers deteriorated from 3.21 to 3.28. The number of ECA countries with changes in their EIU corruption rating is small, as might be expected on a scale with only five categories. Corruption improvements were recorded for Latvia, Lithuania, Slovenia and Turkey, with Poland worsening, in all cases by only one category.

Both of the widely-known composite indexes of corruption show slight improvements for ECA relative to non-ECA countries. Among 152 countries (including most developed nations) with TI index values for both 2002 and 2005, the average ranking among the 28 ECA countries increased from 92 to 88.[10] Of course, evidence from the TI index is at least partly redundant, because it includes the WEF, NIT and EIU measures discussed above. Over the 2002–2005 period, the WBI index ranking for the 28 ECA countries improved from 120 to 110. As with the TI index, evidence from the WBI index should not be interpreted as being fully independent from some of the trends reported above.

Although the various data sources agree with each other – and with the BEEPS evidence – that corruption has tended to decline for ECA overall, there is less agreement about which countries in the region experienced the most improvement. Correlations among the expert-based assessments of changes from 2002 to 2005 are very low and a few are even negative. Poland is a case on which sources generally agreed, with worse ratings in 2005 than in 2002 from NIT, CPIA and EIU. The

Table 4: Changes in World Bank's CPIA Question 16 2002–2004, regional averages

	Change 2002–2004	2002	2004
East Asia/Pacific (18)	+0.32	2.79	3.12
East Europe/Central Asia (27)	+0.19	3.11	3.30
Sub-Saharan Africa (45)	+0.12	2.83	2.95
Latin America/Caribbean (28)	+0.09	3.45	3.54
South Asia (7)	+0.00	3.21	3.21
Middle East/North Africa (9)	−0.06	2.95	2.89
All (134)	+0.14	3.04	3.18

Numbers in parentheses indicate number of countries in each regional group. Table shows regional averages for World Bank's Country Policy and Institutional Assessment rating on corruption, accountability and transparency in the public sector, 2002 and 2004. See Appendix 3 for complete question wording. Regions are listed in order from most to least average progress over time.

Czech Republic is a case of disagreement, with improved ratings in NIT and CPIA despite being downgraded by ICRG.

Contradictory *changes*, however, are potentially consistent with converging views about appropriate ratings *levels*. Changes in expert-based ratings do not always reflect a belief that actual conditions have changed, but often are intended to correct a previous year's rating that in retrospect appears too high or too low. The decrease for Czech Republic by ICRG appears inconsistent with its improvement in NIT and CPIA. But the ICRG in 2002 ranked it higher among ECA countries than the other sources did. Ratings changes for Czech Republic reflect a convergence in assessments among those three sources, and with the WEF and BEEPS data, which also tend to place Czech in the upper half of the ECA rankings, but not among the top 5 or 6. If converging assessments of this sort are more the rule than the exception, one would expect intercorrelations among these expert sources to be higher in 2005 than in 2002. The data confirm this prediction: the mean of the six intercorrelations among ICRG, NIT, CPIA and EIU increases from 0.78 in 2002 to 0.85 in 2005.

Convergence of this sort represents a likely reduction in measurement error. A more fundamental explanation for apparent disagreement among corruption indicators is that they do not all measure exactly the same concept. These differences are easily seen in the definitions in Appendix B. Most notably, the CPIA attempts to measure not only corruption in the public sector, but also 'transparency and accountability'.

4. What aspects of corruption are the broad indicators measuring?

The prevalence and conceptual variety of corruption measures in the 2005 BEEPS and WEF surveys can help identify which aspects of corruption in ECA countries are best captured by broader, perception-based measures – including the BEEPS question on corruption as an obstacle to doing business. Table 5 reports correlations of NIT, ICRG, CPIA, EIU, the BEEPS 'obstacle' question, and the TI and WBI composite indexes with a comprehensive set of corruption measures included in the BEEPS and WEF.

The assessment criteria for NIT, ICRG and CPIA reflect roughly equal mixtures of administrative corruption and state capture, while the extremely brief criteria for EIU, 'how pervasive is corruption by public officials?', is consistent with both types. The correlations with BEEPS variables suggest, however, that all of these sources – particularly the CPIA – are measuring primarily administrative corruption. Among the various BEEPS measures, bribes in business licenses and permits and in tax collection are most strongly correlated with the broadly-defined

corruption measures from other sources. None of the four broad indicators is strongly correlated with bribes for influencing legislation, or with measures of the impacts on business of bribing to affect Parliamentary votes or government decrees.

Correlations of NIT, ICRG, CPIA and EIU with the various WEF firm survey corruption variables show a broadly similar pattern. Each is strongly correlated with bribes for utility connections, exports and

TABLE 5: What aspects of corruption are broad perception indicators measuring? Correlation from the 2005 BEEPS and WEF

	NIT	ICRG	CPIA	EIU	Obstacle	TI	WBI
Business environment and enterprise performance							
Bribe tax (Q40)	0.54	− 0.50	− 0.44	0.52	0.37	− 0.48	**− 0.57**
Bribe frequency (Q39a)	0.57	− 0.52	− 0.54	0.80	0.60	− 0.60	**− 0.67**
Bribe predictability (Q39b)	0.56	− 0.48	− 0.52	0.65	0.52	− 0.61	− 0.65
Obstacle to business (Q54q)	0.31	− 0.35	− 0.12	0.50	—	− 0.46	**− 0.35**
Kickback in govt. contracts (Q42)	0.14	− 0.04	− 0.04	0.19	0.33	− 0.15	− 0.11
Administrative corruption							
Utilities (Q41a)	0.50	− 0.34	− 0.52	0.60	0.54	− 0.60	− 0.59
Licenses & permits (Q41b)	0.68	− 0.58	− 0.66	0.82	0.59	− 0.70	− 0.72
Government contracts (Q41c)	− 0.08	− 0.07	*0.21*	0.01	0.57	− 0.06	− 0.01
Health/safety inspections (Q41d)	0.05	− 0.33	*0.02*	0.38	0.54	− 0.17	− 0.15
Fire/building inspections (Q41e)	0.50	− 0.52	− 0.43	0.58	0.18	− 0.45	− 0.51
Environmental inspections (Q41f)	0.43	− 0.42	− 0.34	0.52	0.53	− 0.40	− 0.51
Taxes & tax collections (Q41g)	0.65	− 0.58	− 0.68	0.79	0.49	− 0.66	− 0.71
Customs/imports (Q41h)	0.38	− 0.41	− 0.35	0.58	0.72	− 0.46	− 0.49
Courts (Q41i)	0.30	− 0.40	− 0.22	0.47	0.67	− 0.41	− 0.41
State capture							
Influence legislation/rules (Q41j)	0.03	− 0.20	*0.03*	0.09	0.62	− 0.11	**− 0.15**
Impact of capture: parliament (Q44a)	0.15	− 0.39	− 0.08	0.36	0.68	− 0.28	− 0.19
Impacted of capture: govt. off. (Q44b)	0.13	− 0.29	− 0.06	0.32	0.73	− 0.27	**− 0.21**
World Economic Forum							
Favoritism in decisions	− 0.34	0.49	0.28	− 0.43	− 0.52	0.52	0.45
Diversion of public funds	− 0.52	0.56	0.42	− 0.66	− 0.83	0.67	**0.63**
Business costs of corruption	− 0.72	0.65	0.72	− 0.79	− 0.68	0.80	**0.81**
Financial honesty of politicians	− 0.22	0.44	0.04	− 0.36	− 0.63	0.40	**0.31**
Irregular payments in . . .							
exports & imports	− 0.73	0.59	0.77	− 0.68	− 0.57	**0.85**	**0.77**
public utilities	− 0.73	0.58	0.78	− 0.73	− 0.55	**0.86**	**0.79**
tax collection	− 0.74	0.69	0.77	− 0.68	− 0.49	**0.82**	**0.80**
public contracts	− 0.26	0.05	0.27	− 0.10	− 0.40	**0.53**	**0.35**
judicial decisions	− 0.69	0.57	0.67	− 0.57	− 0.55	**0.82**	**0.76**

Table entries are correlations of broad corruption indicators (shown in columns) with more specific indicators from BEEPS and WEF firm surveys (shown in rows). Correlations with counterintuitive signs are shown in italics. Correlations in bold indicate correlation by construction; i.e. the variable is a component of the composite index in question.

imports, and tax collection: correlations with a WEF measure of favoritism in decisionmaking are more modest. There is one state capture measure in WEF, however, which is strongly correlated with the four broad indicators: the WEF question on 'business costs of corruption', defined in terms of other firms' illegal payments to influence government laws and policies. With this single exception, data from the two firm surveys indicate that NIT, ICRG, CPIA and EIU are measuring administrative corruption much better than state capture.

A striking finding from the BEEPS and WEF data is the absence of any significant link between corruption in public procurement and the broad, perception-based measures. Of the 12 correlations between three firm-survey variables on bribery in procurement, and the four broad indicators, the highest correlation is 0.27. The third-strongest of these 12 correlations, between CPIA and a BEEPS measure even has a perverse sign.

This weak relationship could be attributable in part to lack of good information, if most firms never sell their products or services to government agencies. Accordingly, we re-calculated the two BEEPS items on corruption in public procurement, deleting the roughly four-fifths of firms in the sample reporting no sales to their government. Correlations with a few of the broad perception-based measures strengthen somewhat, but remain far weaker than any of the other administrative corruption correlations.

Two WEF variables measure business executives' perceptions of other aspects of corruption that pertain more to misappropriation of taxpayer funds by government officials, one entitled 'diversion of public funds', and the other 'public trust in the financial honesty of politicians'. Diversion of funds is most strongly correlated with EIU (0.66), among the four expert-based indicators. It is most weakly correlated (0.42) with CPIA – despite the fact that of the four only the CPIA explicitly includes diversion of funds in its definition. Public trust in the honesty of politicians is correlated at 0.44 with ICRG, but only at 0.04 for CPIA.

The dimensions of corruption that present the largest obstacles to doing business are likely to vary not only across countries, but also across firms within a country. Gray, Hellman and Ryterman (2004) use BEEPS II data to run country-specific, cross-firm regressions of the 'obstacle' measure on several administrative corruption measures (also from the BEEPS), among other regressors. In some countries, they find bribes paid in dealing with courts to be a significant obstacle, while in others bribes paid to obtain business licenses are significant. The fifth column of results in Table 6 is a cruder look into what forms of corruption appear most often to represent a serious obstacle, using the BEEPS III data. Unlike Gray, Hellman and Ryterman (2004), it does not disaggregate by

country, or control for other variables. With those caveats, the broad perceptions measure of corruption in BEEPS – the 'obstacle' measure – is found to be correlated more strongly, on average, with the state capture questions in BEEPS, and to a lesser extent in WEF, than with their administrative corruption questions. These findings are in stark contrast to those reported above for NIT, ICRG, CPIA and EIU. Among all of the WEF and BEEPS indicators, the 'obstacle' variable is most highly correlated (0.83) with the WEF 'diversion of public funds'. This result is somewhat surprising, as firms are not well-placed to have first-hand knowledge on diversion of public funds, in contrast to their frontline position with respect to many forms of administrative corruption, state capture and procurement fraud.

The EIU and ICRG indicators are produced by commercial firms specializing in assessing risk to overseas investors. They might therefore focus on assessing corruption conditions faced by foreign-owned companies, which may sometimes differ from those faced by domestically owned firms. Accordingly, we re-calculated all of the country-level BEEPS corruption measures using only the 12 per cent of firms that were majority foreign-owned. On average, the EIU indicator is no more highly correlated with these BEEPS measures than with those calculated using all firms. If conditions facing foreign owned firms are different, the EIU does not appear to measure those differences effectively. Most correlations between BEEPS questions and ICRG, however, are higher (by 0.05 on average, in absolute value) when BEEPS measures are calculated only for firms that are majority foreign-owned. The correlations with ICRG strengthen the most for BEEPS questions on bribes paid for utility connections, and for environmental, health and safety inspections. Bribe *frequency* for foreign-owned firms is no different on average than for other firms, but the average bribe *tax* they report is slightly lower.

The correlations of broadly-defined corruption measures with more specific questions in the BEEPS and WEF firm surveys described above can help reveal what information underlies subjective judgments regarding corruption in ECA countries. Correlations of the composite indexes with BEEPS and WEF, reported in the final two columns of Table 5, must be interpreted differently, because many of the firm survey variables are correlated by construction with the composite indexes. The WEF administrative corruption variables (the bottom five in Table 5) are part of the TI index. Not surprisingly, they are more strongly correlated with TI than are the other WEF variables, and the BEEPS variables, which are not components of TI.

Because the WEF administrative corruption variables enter the TI index as three separate sources – for the three most recent annual surveys – the overall TI index is likely to emphasize administrative

corruption more than state capture. Therefore, we can expect the administrative corruption measures in BEEPS to be more strongly correlated than the BEEPS state capture measures with TI, even though BEEPS is not a component of the TI index. That is in fact what is found in Table 5. Correlations of TI with the three state capture variables in BEEPS range from 0.11 to 0.28. Correlations with TI exceed 0.40 for 7 of the 9 BEEPS administrative corruption measures, with the highest for tax collection (0.66) and business licenses (0.70).

For WBI – in marked contrast to TI – variables included in the index are no more highly correlated with it than are variables excluded from the index. The two BEEPS variables most highly correlated with WBI and with TI are bribe frequency in the areas of business licenses and permits, and in tax collection. Neither of these variables is a component of the WBI index. State capture measures in BEEPS are even more weakly correlated with WBI than with TI, although some of them are components of WBI (unlike the case for TI). Also, BEEPS receives only an extraordinarily small weight in the WBI index, only about 1/18 of the weight given to NIT. Due to the huge weight for NIT, the correlation of WBI with NIT is 0.97; it is not surprising that results in Table 5 for WBI closely mirror those for NIT, but with the signs reversed.

Corruption in public procurement, as measured by two BEEPS questions, has a near zero correlation with the TI and WBI indexes. The WEF question on 'irregular payments' needed to obtain public contracts is a component of both the TI and WBI indexes, so it is moderately correlated with them. But correlations of both indexes are far higher for the other four WEF 'irregular payments' questions related to administrative corruption. These may be the most noteworthy findings from this exercise, as graft in public procurement receives more publicity than any other aspects of corruption. Media reports on procurement fraud in a country are often accompanied by references to its TI ranking. However, corruption in procurement – as reported in firm surveys – has little to do with rankings on TI and WBI.

Factor analysis is an alternative approach for analyzing the content of broad corruption indicators. A factor analysis of 12 BEEPS variables (all but the first five listed in Table 5) yields two significant factors (with eigenvalues exceeding one) that together explaining 85 per cent of the variation in the data. One of these is clearly identifiable as a 'state capture' factor: the three variables on unofficial payments to influence legislation and rules all load most heavily on it (Q41j, Q44a, Q44b). The second factor reflects administrative corruption: variables loading most heavily on it include payments to obtain business licenses (Q41b), deal with fire and building inspections (Q41e), and deal with taxes and tax collection (Q41g). Adding one of the more broadly-defined perceptions

measures to these 13 in a factor analysis, we can observe which of these factors it loads most heavily on, and infer which of these two major types of corruption it is best capturing.

When NIT is added in the factor analysis, it has a large positive loading on the administrative corruption factor, but a small negative loading on the state capture factor. Very similar results are found if any of the other broadly-defined corruption indicators listed across the columns of Table 5 is substituted for NIT.

Factor analyses based on the 9 WEF variables in Table 5 produce similar findings to those based on BEEPS variables. Two significant factors explain 95 per cent of the variation. These factors are again clearly identifiable as state capture and administrative corruption, with the first four WEF variables listed in Table 5 loading most heavily on one factor, and the other five loading most heavily on the other. Corruption measures from NIT, CPIA, TI and WBI all load mostly on the administrative corruption factor when any one of them is added as a 10th variable. The ICRG, EIU and the 'obstacle' measure from BEEPS load more equally across the two factors, although still somewhat more strongly on administrative corruption than on state capture.

The weak link between state capture measures in BEEPS and in broader perceptions-based indicators is due in part to Belarus and Uzbekistan. These two countries are 'outliers' in being rated lower by other sources than by most of the BEEPS questions, particularly those on state capture. The first *Anticorruption in Transition* report (World Bank 2000) attributed low levels of state capture in Belarus and Uzbekistan to their relatively small private sectors and 'the continued existence of authoritarian controls'. The third report (Gray and Anderson 2006) discusses in greater detail the possibility that corruption takes different forms, not easily measured by firm surveys, in autocratic regimes. Indeed, there is no necessary contradiction between infrequent bribery of public officials by firms (i.e., relatively good performance on BEEPS) on the one hand, and excessive state involvement in the economy, absence of protection for whistleblowers and journalists, etc. on the other (i.e., a poor rating on NIT; see Appendix 2 for criteria).

Omitting Belarus and Uzbekistan, correlations of NIT and CPIA with the BEEPS state capture measures (and with its 'obstacle' measure) are somewhat stronger than those reported in Table 5. Belarus is not covered by EIU, and Uzbekistan is not covered by ICRG, and results for those sources change by less when those countries are deleted. Neither Belarus nor Uzbekistan is included in the WEF sample, so none of the correlations reported in the lower part of Table 5 are affected by their deletion. The factor analysis using WEF data also indicated that the broadly-defined corruption indicators were mostly measuring administrative

corruption rather than state capture. Even without these two countries, therefore, the evidence indicates that the broad measures reflect administrative corruption much more than they reflect state capture.

5. Conclusions

The BEEPS and other sources of corruption data indicate that corruption in ECA overall is declining. Gray, Hellman and Ryterman (2000: 40) attribute part of the decline in corruption measured in BEEPS I and II (1999 to 2002) to optimism, perhaps associated with relatively strong economic performance. Continued favorable economic conditions may similarly play some role in the improvements measured by BEEPS II and III (2002 to 2005. Expert ratings can also be affected by recent economic performance: other things equal one might infer that corruption must not be too severe if growth is strong. The Bangladeshi case suggests that such inferences are not paramount in making assessments; it routinely ranks at the bottom of the TI index, despite experiencing fairly rapid growth in recent years. For small countries, however, on which experts tend to have less information, corruption assessments may rely more heavily on proxies such as economic conditions, or type of political regime. Rose and Mishler (2007) show that public perceptions of corruption in Russia appear to be driven more by media reports than by actual experiences. More research is needed concerning the impact of optimism, recent economic performance, and highly publicized corruption scandals on country-level corruption indicators of the expert-assessment type as well as firm and household surveys.

More inquiry is needed into the actual content of commonly-used indicators, as distinct from their purported content. The criteria for several sources (including ICRG, CPIA, and NIT) place great weight on state capture, but appear at least for the ECA region to be measuring primarily administrative corruption.

Conceptual, methodological and empirical materials strongly support the message that no single corruption measure, nor single data source on corruption, is most appropriate for all purposes. Expert ratings are defined too vaguely and broadly-and constructed too non-transparently – to be suitable for some purposes. For example, it is difficult to hold governments responsible for improving their scores on such measures, as a condition for receiving aid, if there is little indication of how scores can be improved.

Composite indexes of corruption should be used with more caution by development agencies and by researchers. There should be more examination of the criteria and methods of their underlying sources and their degree of interdependence. Depending on one's purposes,

customized indexes based on a subset of the TI or WBI components might be more appropriate. If the underlying data were made more accessible, data users could choose the weights they deem appropriate for their purposes, in customizing an index. They could also compare two countries, or two time periods within a country, using only data sources common to both. All users of the composite indexes and their perceptions-based components should follow TI's example and acknowledge that these are measures of corruption *perceptions*, not of corruption.

In comparison to broad expert assessments, a virtue of BEEPS (and WEF) is 'unbundling' corruption into a large set of survey questions. Firm-level analyses, e.g. on firm characteristics associated with different forms of corruption, can be conducted using BEEPS. An important limitation of firm surveys such as BEEPS however is that they 'provide a very incomplete measurement of corruption' (Gray, Hellman, and Ryterman 2004: 54) by measuring only interactions between firms and public officials.

To improve on the existing set of country-level corruption indicators, more data collection is needed on several fronts. First, the BEEPS should be replicated for other regions. The World Bank, in partnership with some of the regional development banks, is already working towards this goal. Second, firm surveys should be complemented by more systematic household surveys measuring experiences with corruption and other governance problems. Transparency International's 'Global Corruption Barometer' is a promising development in this regard, but conducting nationally-representative surveys of households remains a severe challenge in many developing countries. Third, public officials surveys (sporadically conducted by the Bank in a small number of countries) should be standardized and scaled up, with a focus on assessing aspects of public sector corruption and other governance deficiencies not manifested in either state-enterprise or state-household transactions. Finally, existing efforts to collect data on laws and practices intended to prevent corruption should be scaled up, to provide more 'actionable' indicators appropriate for monitoring reform commitment and progress. Promising developments in this area include the Public Integrity Index, the International Budget Project, and the Public Expenditure and Financial Accountability (PEFA) indicators.[11]

NOTES

1. Information and data for the World Bank's investment climate assessment surveys are available at http://iresearch.worldbank.org/ics/jsp/index.jsp.
2. See http://www.uneca.org/agr/.
3. See http://www.odi.org.uk/WGA_Governance/Index.html.
4. For example, the lowest rating of 7 implies an 'absence of practices that adhere to basic human rights standards, democratic norms, and the rule of law' on the NIT corruption index and on its

other six indexes: National Democratic Governance, Electoral Process, Civil Society, Independent Media, Local Democratic Governance, and Judicial Framework and Independence.

5. This information was provided by Jim Anderson, who closely examined data sources in the TI and WBI indexes.
6. The WEF (www.weforum.org) publishes an annual 'Global Competitiveness Report' and the IMD (http://www.imd.ch/research/publications/wcy/index.cfm) publishes the 'World Competitiveness Yearbook'. Both organizations list their partner institutes on their web sites.
7. The TI index uses the most recent three years of data for WEF and IMD.
8. The ICRG's current editor (like representatives for EIU and Global Insight) readily acknowledges looking at the TI ratings as one of many sources of information for the ratings. He did not know however whether realignment with TI explained the data break in 2001, which occurred under his predecessor.
9. The 2004 ratings were produced in late 2004 and early 2005, so provide a better comparison than the 2005 ratings (finalized in early 2006) with the 2005 BEEPS. The 2002 CPIA ratings were produced in mid and late 2002, so provide the best match with the 2002 BEEPS.
10. The TI index provided on TI's official web site lists somewhat fewer countries, namely only those for which at least three data sources were available. Johann Lambsdorff lists additional countries for which only one or two data sources were available, on the web site of the Internet Center for Corruption Research (http://www.icgg.org/).
11. Information on these initiatives can be found, respectively, at http://www.globalintegrity.org/, http://www.internationalbudget.org/, and http://www.pefa.org/.

REFERENCES

Anderson, J. and Gray, C. (2006) *Anticorruption in Transition 3: Who is Succeeding, and Why?* Washington DC: The World Bank.

Arndt, C. and Oman, C. (2006) *Uses and Abuses of Governance Indicators*. Paris: OECD Development Centre.

Economic Commission for Africa (2005). *Striving for Good Governance in Africa*. Addis Ababa: UNECA.

Gray, C.; Hellman, J. and Ryterman, R. (2004) *Anticorruption in Transition 2: Corruption in Enterprise-State Interactions in Europe and Central Asia 1999–2002*. Washington DC: The World Bank.

Hyden, G.; Court, J. and Mease, K. (2004) *Making Sense of Governance: Empirical Evidence from Sixteen Developing Countries*. Boulder, Co: Lynne Rienner.

IMD (2005). World Competitiveness Yearbook 2005. Lausanne, Switzerland: IMD. (http://www01.imd.ch/wcc/ranking/).

Kaufmann, D. (2005) 'Myths and Realities of Governance and Corruption', in A. Lopez- Claros, M.E. Porter and K. Schwab (eds.) *Global Competitiveness Report 2005–2006: Policies Underpinning Rising Prosperity*. Houndmills, UK: Palgrave Macmillan for the World Economic Forum.

Kaufmann, D.; Kraay, A. and Mastruzzi, M. (2006) 'Governance Matters V: Governance Indicators for 1996–2005'.' World Bank Policy Research Working Paper 4012.

Kaufmann, D.; Kraay, A. and Zoido-Lobaton, P. (2000) 'Governance Matters: From Measurement to Action', *Finance and Development*, 37, 2, 10–13.

Lambsdorff, J.G. (undated) 'Measuring the Dark Side of Human Nature: The Birth of the Corruption Perceptions Index' (http://www.icgg.org/corruption.cpi_childhooddays.html).

Lambsdorff, J.G. (2005a) 'The Methodology of the 2005 Corruption Perceptions Index' (http://www.icgg.org/downloads/CPI_Methodology.pdf).

Lambsdorff, J.G. (2005b) 'Determining Trends for Perceived Levels of Corruption.' University of Passau Discussion Paper V-38-05.

Lopez-Claros, A.; Porter, M.E. and K. Schwab (2005) *Global Competitiveness Report 2005–2006: Policies Underpinning Rising Prosperity*. Houndmills, UK: Palgrave Macmillan for the World Economic Forum.

Mauro, P. (1995) 'Corruption and Growth', *Quarterly Journal of Economics*, 110, 3, 681–712.

PRS Group (2003) 'A Brief Guide to the Ratings System' [for the International Country Risk Guide]. Syracuse, NY: PRS Group.

Rose, R. and Mishler, W. (2007) 'Bridging the Gap Between Experience and Perception of Corruption', Studies in Public Policy No. 432. Aberdeen: University of Aberdeen.

Sandholtz, W. and Koetzle, W. (2000) 'Accounting for Corruption: Economic Structure, Democracy and Trade', *International Studies Quarterly*, 44, 31–50.

Surowiecki, J. (2004) *The Wisdom of Crowds*. New York: Doubleday.

Swamy, A., Knack, S., Lee, Y. and Azfar, O. (2001) 'Gender and Corruption', *Journal of Development Economics*, 64, February 2001, 25–55.
Economic Commission for Africa (2005) *Striving for Good Governance in Africa: Synopsis of the 2005 AGR.* Addis Ababa: UNECA.
World Bank (2000) *Anticorruption in Transition: a Contribution to the Policy Debate.* Washington DC: World Bank.

APPENDIX: DEFINITIONS OF CORRUPTION INDICATORS

1. Survey Questions from BEEPS (Business Environment and Enterprise Performance Survey)

Bribe frequency & predictability: Thinking about officials, would you say the following statements are always, usually, frequently, sometimes, seldom or never true? Never = 1, seldom = 2, sometimes = 3, frequently = 4, usually = 5, always = 6

'It is common for firms in my line of business to have to pay some irregular 'additional payments/gifts' to get things done with regard to customs, taxes, licenses, regulations, services etc.' (Q39a)

'Firms in my line of business usually know in advance about how much this 'additional payment/gift is.' (Q39b)

Bribe tax (Q40): On average, what per cent of total annual sales do firms like yours typically pay in unofficial payments/gifts to public officials?

Corruption as a problem doing business (Q54q): Can you tell me how problematic are these different factors for the operation and growth of your business: . . . Corruption (No obstacle = 1 Minor obstacle = 2 Moderate obstacle = 3 Major obstacle = 4)

Kickback for government contracts (Q42): When firms in your industry do business with the government, what per cent of the contract value would be typically paid in additional or unofficial payments/gifts to secure the contract? . . .%

Sector-specific bribe frequency (Q41): Thinking now of unofficial payments/gifts that a firm like yours would make in a given year, could you please tell me how often would they make payments/gifts for the following purposes: (Never = 1, seldom = 2, sometimes = 3, frequently = 4, usually = 5, always = 6)

To get connected to and maintain public services (electricity and telephone) (Q41a)
To obtain business licenses and permits (Q41b)
To obtain government contracts (Q41c)
To deal with occupational health and safety inspection (Q41d)
To deal with fire and building inspections (Q41e)
To deal with environmental inspections (Q41f)
To deal with taxes and tax collection (Q41g)
To deal with customs/imports (Q41h)

To deal with courts (Q41i)

To influence the content of new legislation rules decrees etc. (Q41j)

Impact of capture (Q44): It is often said that firms make unofficial payments/gifts, private payments or other benefits to public officials to gain advantages in the drafting of laws, decrees, regulations, and other binding government decisions. To what extent have the following practices had a direct impact on your business? (No impact, minor impact, moderate impact, major impact, decisive impact)

Private payments/gifts or other benefits to Parliamentarians to affect their vote (Q44a)

Private payments/gifts or other benefits to Government officials to affect the content of government decrees (Q44b)

2. Nations in Transit (http://www.freedomhouse.org/research/nattransit.htm)

For all 28 countries and territories in Nations in Transit 2005, Freedom House, in consultation with the report authors and a panel of academic advisers, has provided numerical ratings [on corruption and six other variables]. The ratings are based on a scale of 1 to 7, with 1 representing the highest and 7 the lowest level.

The ratings follow a quarter-point scale. Minor to moderate developments typically warrant a positive or negative change of a quarter (0.25) to a half (0.50) point. Significant developments typically warrant a positive or negative change of three-quarters (0.75) to a full (1.00) point. It is rare that the rating in any category will fluctuate by more than a full point (1.00) in a single year.

The ratings process for Nations in Transit 2005 involved four steps:

(1) Authors of individual country reports suggested preliminary ratings in all seven categories covered by the study.

(2) The US and CEE-NIS (Central and Eastern Europe-Newly Independent States) academic advisers evaluated the ratings and made revisions.

(3) Report authors were given the opportunity to dispute any revised rating that differed from the original by more than 0.50 point.

(4) Freedom House refereed any disputed ratings and, if the evidence warranted, considered further adjustments. Final editorial authority for the ratings rested with Freedom House.

Corruption. Ratings reflect public perceptions of corruption, the business interests of top policy makers, laws on financial disclosure and conflict of interest, and the efficacy of anticorruption initiatives.

(1) Has the government implemented effective anticorruption initiatives?

(2) Is the country's economy free of excessive statement involvement?
(3) Is the government free from excessive bureaucratic regulations, registration requirements, and other controls that increase opportunities for corruption?
(4) Are there significant limitations on the participation of government officials in economic life?
(5) Are there adequate laws requiring financial disclosure and disallowing conflict of interest?
(6) Does the government advertise jobs and contracts?
(7) Does the state enforce an effective legislative or administrative process – particularly one that is free of prejudice against one's political opponents – to prevent, investigate, and prosecute the corruption of government t officials and civil servants?
(8) Do whistle-blowers, anticorruption activists, investigators, and journalists enjoy legal protections that make them feel secure about reporting cases of bribery and corruption?
(9) Are allegations of corruption given wide and extensive airing in the media?
(10) Does the public display a high intolerance for official corruption?

3. World Bank Country Policy and Institutional Assessment (CPIA)

Transparency, Accountability, and Corruption in the Public Sector

This criterion assesses the extent to which the executive can be held accountable for its use of funds and the results of its actions by the electorate and by the legislature and judiciary, and the extent to which public employees within the executive are required to account for the use of resources, administrative decisions, and results obtained. Both levels of accountability are enhanced by transparency in decision-making, public audit institutions, access to relevant and timely information, and public and media scrutiny. A high degree of accountability and transparency discourages corruption, or the abuse of public office for private gain. National and sub-national governments should be appropriately weighted. Each of three dimensions should be rated separately: (a) the accountability of the executive to oversight institutions and of public employees for their performance; (b) access of civil society to information on public affairs; and (c) state capture by narrow vested interests. For the overall rating, these three dimensions should receive equal weighting. A rating for each dimension should be provided in the write-up along with its justification. Ratings range from 1 (lowest) to 6 (highest). The table below describes the criteria associated with each ratings level.

1a) There are no checks and balances on executive power. Public officials use their positions for personal gain and take bribes openly. Seats in the legislature and positions in the civil service are often bought and sold.

1b) Government decision-making is secretive. The public is prevented from participating in or learning about decisions and their implications.

1c) The state has been captured by narrow interests (economic, political, ethnic, and/or military). Administrative corruption is rampant.

2a) There are only ineffective audits and other checks and balances on executive power. Public officials are not sanctioned for failures in service delivery or for receiving bribes.

2b) Decision making is not transparent, and government withholds information needed by the public and civil society organizations to judge its performance. The media are not independent of government or powerful business interests.

2c) Boundaries between the public and private sector are ill-defined, and conflicts of interest abound. Laws and policies are biased towards narrow private interests. Implementation of laws and policies is distorted by corruption, and resources budgeted for public services are diverted to private gain.

3a) External accountability mechanisms such as inspector-general, ombudsman, or independent audit may exist, but have inadequate resources or authority.

3b) Decision making is generally not transparent, and public dissemination of information on government policies and outcomes is a low priority. Restrictions on the media limit its potential for information-gathering and scrutiny.

3c) Elected and other public officials often have private interests that conflict with their professional duties.

4a) External accountability mechanisms limit somewhat the degree to which special interests can divert resources or influence policy making through illicit and non-transparent means. Risks and opportunities for corruption within the executive are reduced through adequate monitoring and reporting lines.

4b) Decision making is generally transparent. Government actively attempts to distribute relevant information to the public, although capacity may be a constraint. Significant parts of the media operate outside the influence of government or powerful business interests, and media publicity provides some deterrent against unethical behavior.

4c) Conflict of interest and ethics rules exist and the prospect of

sanctions has some effect on the extent to which public officials shape policies to further their own private interests.

5a) Accountability for decisions is ensured through a strong public service ethic reinforced by audits, inspections, and adverse publicity for performance failures. The judiciary is impartial and independent of other branches of government. Authorities monitor the prevalence of corruption and implement sanctions transparently.

5b) The reasons for decisions, and their results and costs, are clear and communicated to the general public. Citizens can obtain government documents at nominal cost. Both state-owned (if any) and private media are independent of government influence and fulfill critical oversight roles.

5c) Conflict of interest and ethics rules for public servants are observed and enforced. Top government officials are required to disclose income and assets, and are not immune from prosecution under the law for malfeasance.

6 Criteria for '5' on all three sub-ratings are fully met. There are no warning signs of possible deterioration, and there is widespread expectation of continued strong or improving performance.

4. International Country Risk Guide (http://www.prsgroup.com/icrg/icrg.html)

Corruption. This is an assessment of corruption within the political system. Such corruption is a threat to foreign investment for several reasons: it distorts the economic and financial environment; it reduces the efficiency of government and business by enabling people to assume positions of power through patronage rather than ability, and, last but not least, introduces an inherent instability into the political process.

The most common form of corruption met directly by business is financial corruption in the form of demands for special payments and bribes connected with import and export licenses, exchange controls, tax assessments, police protection, or loans. Such corruption can make it difficult to conduct business effectively, and in some cases may force the withdrawal or withholding of an investment.

Although our measure takes such corruption into account, it is more concerned with actual or potential corruption in the form of excessive patronage, nepotism, job reservations, 'favor-for-favors', secret party funding, and suspiciously close ties between politics and business. In our view these insidious sorts of corruption are potentially of much greater risk to foreign business in that they can lead to popular discontent, unrealistic and inefficient controls on the state economy, and encourage the development of the black market.

The greatest risk in such corruption is that at some time it will become so overweening, or some major scandal will be suddenly revealed, as to provoke a popular backlash, resulting in a fall or overthrow of the government, a major reorganizing or restructuring of the country's political institutions, or, at worst, a breakdown in law and order, rendering the country ungovernable.

5. World Economic Forum (http://www.weforum.org/)

Executive Opinion Survey: Corruption-related questions, included in both 2002–3 and 2005–6 surveys
Irregular payments in exports and imports: In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with export and import permits? (1 = common, 7 = never occurs)
Irregular payments public utilities: In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes when getting connected to public utilities? (1 = common, 7 = never occurs)
Irregular payments in tax collection: In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with annual tax payments? (1 = common, 7 = never occurs)
Irregular payments in public contracts: In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with public contracts (investment projects)? (1 = common, 7 = never occurs)
Irregular payments in judicial decisions: In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with getting favorable judicial decisions? (1 = common, 7 = never occurs)
Business costs of corruption: Do other firms' illegal payments to influence government policies, laws or regulations impose costs or otherwise negatively affect your firm? (1 = impose large costs, 7 = impose no costs/ not relevant)
Favoritism in decisions of government officials: When deciding upon policies and contracts, government officials (1 = usually favor well-connected firms and individuals, 7 = are neutral among firms and individuals)
Diversion of public funds: In your country, diversion of public funds to companies, individuals or groups due to corruption (1 = is common, 7 = never occurs)

Public trust of politicians: Public trust in the financial honesty of politicians is (1 = very low, 7 = very high)

STEPHEN KNACK
Lead Economist,
World Bank
1818 H St. NW, Washington DC 20433
e-mail: sknack@worldbank.org