

WPS3968

Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators

Stephen Knack*

World Bank Policy Research Working Paper 3968, July 2006

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

*Address correspondence to Stephen Knack, Senior Research Economist, World Bank, 1818 H Street, N.W., Washington, D.C. 20433. Email: sknack@worldbank.org. Phone: 202-458-9712. Fax: 202-522-1154.

This paper benefited from enormously valuable suggestions and comments from James Anderson and Cheryl Gray.

Abstract

This paper assesses corruption levels and trends among countries in the transition countries of Eastern Europe and Central Asia (ECA), based on data from several sources that are both widely used and cover most or all countries in the region. Data from firm surveys tend to show improvement in most types of administrative corruption, but little change in “state capture” in the region. Broader, subjective corruption indicators tend to show somewhat greater improvement in ECA than in non-ECA countries on average. A “primer on corruption indicators” discusses definitional and methodological differences among data sources that may account in large part for the apparently conflicting messages they often provide. This discussion concludes that depending on one’s purpose, it may be more appropriate to use data from a single source rather than a composite index, because of the loss of conceptual precision in aggregation. A second conclusion is that the gains in statistical “precision” from aggregating sources of corruption data likely are far more modest than often claimed, because of interdependence among data sources. The range of detailed corruption measures available in firm surveys are exploited to show that broad, perceptions-based corruption assessments appear to measure primarily administrative corruption, despite their stated criteria placing great weight on “state capture.” Finally, the paper emphasizes the need for scaling up data initiatives to fill significant gaps between our conceptual definitions of corruption and the operational definition embodied in the existing measures.

1. Introduction

This paper assesses corruption levels and trends among countries in the Eastern Europe and Central Asia (ECA) region, based on data from several sources that are both widely used and cover most or all countries in the region. To make sense of these data, the paper also examines the properties of the corruption indicators themselves. The ECA region is the most appropriate one for this analysis as it has the richest set of available country-level data on corruption. Because the various data sources are not always in agreement, we examine definitional and methodological issues in a “primer on corruption indicators” before reporting evidence on levels and trends.

Section 2 discusses definitional and methodological differences among data sources that may account in large part for the apparently conflicting messages they often provide. Section 3 assesses claims that the solution to disagreement among sources is to aggregate them, on the assumption that they collectively are more informative than is any single source. Section 4 presents evidence from the various data sources on corruption trends in the ECA region. The range of corruption measures available in firm surveys are exploited in section 5 to show that broad, perceptions-based corruption assessments appear to measure primarily administrative corruption. Section 6 concludes with recommendations. The main points of the paper are summarized below for convenience:

- Existing corruption indicators differ importantly in the aspects of corruption they purport to measure, in clarity and breadth of definition, and in the methods and transparency of their assessments. For these reasons, *no one indicator or data source is best for all purposes.*
- Aggregating corruption indicators from numerous sources – with the goal of increasing precision in measurement - does not always produce a more appropriate measure than using a single indicator or data source. One cost that should be considered is the loss of *conceptual* precision through aggregation.

- Gains in statistical precision from aggregating sources of corruption data likely are far more modest than often claimed, because the assumption of independent error in measurement among data sources is violated.
- For various reasons, changes over time in corruption ratings should be interpreted with extreme caution. For example, changes in perceptions indicators from one year to the next often are intended to correct ratings regarded in hindsight as incorrect.
- Enterprise surveys such as BEEPS measure only corrupt transactions between public officials and business firms, and in that sense provide a more limited picture than more broadly-defined corruption measures. The advantage of firm surveys such as BEEPS is in providing narrow, specific indicators such as bribes paid in tax collection or in business licensing, and in providing objective measures on share of firm revenues or contract values paid as bribes to public officials. The BEEPS also allows firm-level analyses, e.g. on which types of firms pay more in bribes.
- Changes over time in corruption levels as measured by firm surveys can produce valid inferences if the survey questions and sample design are identical in both periods, and if other factors are controlled for where necessary. For example, perceptions that corruption is an obstacle to doing business are potentially affected by optimism, or by prevailing economic conditions.
- Data from BEEPS, as well as from the WEF executive opinion surveys, show improvement in the region in most but not all types of administrative corruption between 2002 and 2005, and little change in “state capture.”
- Most of the broad, perceptions indicators of corruption show somewhat greater improvement in ECA than in non-ECA countries on average between 2002 and 2005.
- Sources – including firm surveys - disagree markedly on which ECA countries have improved and which have not. This apparent disagreement is in part attributable to the fact that changes in expert assessments often reflect corrections rather than a belief that corruption has actually improved or worsened. In general, the rankings provided by different sources show convergence between 2002 and 2005, so in that sense are not really inconsistent.
- Detailed corruption questions in BEEPS and WEF can be used to shed light on what aspects of corruption are emphasized by the broad, perceptions-based indicators. The latter appear to measure primarily administrative corruption, rather than “state capture,” and appear to measure corruption in public procurement particularly poorly.
- More research is needed to understand better the informational content and possible biases in existing corruption indicators.

- There is a need to develop more “actionable” indicators, assessed for most developing countries, of public sector policies and institutions potentially important in combating corruption. Monitoring progress on these indicators could provide greater incentive for reform.

2. Properties of Corruption Indicators

There are numerous definitions of corruption in the academic literature and among donor agencies. Most of these definitions are quite broad, and often somewhat vague. Transparency International’s definition, “the misuse of entrusted power for private gain,” is representative.¹ Often, the term “misuse” or “abuse” is further defined to apply only to illegal actions. Accepting the brief conceptual definition offered by TI, “corruption” can be disaggregated along many dimensions:

- By level of political system (central government, provincial, municipal), roughly corresponding to the terms “petty” and “grand” corruption;
- By purpose of the improper actions: to influence the content of laws and rules (“state capture”) or to influence their implementation (“administrative corruption”);
- By the actors involved in the corrupt transaction: various combinations of firms, households, and public officials;
- By characteristics of a particular set of actors, for example bribes required for large v. small firms, or for rich v. poor households;
- By administrative agency or service: tax and customs, business licenses, inspections, utility connections, courts, or public education and health facilities.
- By incidence or magnitude of bribes, or by the uncertainty they create for businesses and households.

Regardless of one’s preferred conceptual definition, the choice of measurement techniques from a limited set of feasible alternatives inevitably produces an implicit

¹ <http://www.transparency.org/faqs/faq-corruption.html>. Also see the definitional discussion in Sandholtz and Koetzle (2000), and sources cited therein.

definition that may 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" for federal countries such as the United States or India?² These sorts of questions typically are not explicitly answered in the methodology of existing country-level indicators. 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,

² The impact of corruption should be lessened if firms and households in corrupt cities or states (e.g. Louisiana or Bihar) can readily move to less corrupt areas (e.g. Minnesota or Kerala).

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 every three years as with BEEPS.

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 B5) from this survey are published for 117 countries in WEF's annual Global Competitiveness Report (Lopez-Claros, Porter and Schwab, 2005). Ratings are computed as the simple average of all executives' responses. Another organization, the IMD, uses a nearly identical methodology, but using 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 it 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.³

Household surveys addressing corruption issues are not quite so well developed as firm surveys. Beginning in 2003, Transparency International annually has sponsored the Global Corruption Barometer (GCB), conducted with assistance from Gallup International's survey network. The number of countries covered has expanded from 44 in 2003, to 64 in 2004, and 69 in 2005. The questions changed almost entirely from 2003 to 2004, but much of the 2004 content remained in the 2005 survey. 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. Most of these household surveys suffer from greater comparability problems than does the BEEPS. For example, the surveys administered by Gallup International (including the GCB) cover only urban households in many countries. Unlike the BEEPS, results from the WVS and regional "Barometer" surveys are made public only with long lags, limiting their value in diagnosing problems and designing policy responses.

Expert assessments of corruption have been most widely used for comparisons across countries and over time. A large and growing number of organizations provide such assessments. 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 informed by a

³ Information and data for the World Bank's investment climate assessment surveys are available at <http://iresearch.worldbank.org/ics/jsp/index.jsp>.

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. Two examples are the UNECA's Africa Governance Indicators (Economic Commission for Africa, 2005) and the World Governance Assessments (Hyden, Court and Mease, 2004). The Africa Governance Indicators (AGI), covering corruption and other governance issues, are based on surveys of elites in 28 countries, conducted in 2002-2003. The AGI "expert panels" varied in size from about 70 to 120 across countries. From 83 total questions in the survey, responses to 7 were used in constructing a "Corruption Control" index for each country.⁴

World Governance Assessments were conducted in late 2000 and early 2001 in 22 developing countries from various regions, including Bulgaria, Kyrgyzstan and Russia from the ECA region. In each country, 35 "well-informed persons" were asked 30 questions, including 3 pertaining to corruption (in business licensing, in the judiciary, and favoritism in applying regulations). Data in 6 of the 22 countries were deemed to be of unacceptably low quality, so the publicly available data set covers only 16 countries.⁵ At this time it is unclear whether or not the World Governance Assessments and the Africa Governance Indicators will fulfill their original intentions to expand their country coverage, and to track changes over time.

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. Moreover,

⁴ See <http://www.uneca.org/agr/>.

⁵ See http://www.odi.org.uk/WGA_Governance/Index.html.

respondents in firm surveys can be asked more specific and objective questions, because they comprise a more homogeneous group. A survey of elites that includes public officials, academics, journalists, etc. must frame questions in such a way that they can be answered meaningfully by all of them, which necessitates broader questions. All of the 83 questions in the UNECA survey, and all 30 in the WGA survey, are subjective, with a standard set of five qualitative response categories.

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 .98 with those proposed by the regions.

A second way in which expert assessments differ from each other is in the extent of documentation they provide regarding definitions and methods. For example, NIT 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. Neither the ratings nor the justifications are publicly released however.⁶ Sources that are more transparent and accountability, as reflected by the availability of detailed assessment criteria and justifications for ratings of each country,

⁶ Ratings, but not justifications, for the IDA-eligible countries will be publicly released for the first time in summer 2006.

arguably will tend to be more accurate in their assessments. At a minimum, one can debate meaningfully the appropriateness of the rating and the validity of the methods and information underlying them. Where definitions are brief, vague, and broad, and ratings are not accompanied by justifications for each country, such debate is impossible.

Corruption indicators also differ in attempting to assess either:

- a) the relative incidence of corrupt transactions, or
- b) the impact of corruption on business, or
- c) the existence of government and other 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 B2 and B4 provide the criteria used by ICRG and NIT. The World Bank's CPIA question 16 (see appendix B3) 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 of corruption indicators may have varying 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 for-profit 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 IDA allocations for the World Bank's lower-income countries, the Bank's country teams could benefit from proposing higher-than-warranted ratings. Country teams may also find their working relations with country counterparts impaired if their assessments are unfavorable. However, statistical analysis finds no evidence that IDA countries are overrated relative to non-IDA countries.⁷

Corruption indicators differ in conceptual breadth; some are more multi-dimensional than others. The ICRG, NIT, and CPIA each provide a single measure of corruption, but one intended to reflect a mix of various aspects of corruption.⁸ 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

⁷ Specifically, if the CPIA corruption ratings are regressed on other available corruption indicators and on an IDA dummy, the coefficient for the latter is negative, instead of positive as implied by the potential incentive bias.

⁸ Question 16 in the CPIA actually contains three sub-ratings, but each of these in turn assesses multiple aspects of corruption (or mechanisms intended to deter it).

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 - less obviously - 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, consider the case of constructing a multi-item index from several of the BEEPS (or WEF) corruption measures. Aggregation of them implies a reduction in conceptual precision. But in this case there is no increase in uncertainty over what is being measured, because 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. Even for NIT with its detailed criteria, is each of the 10 criteria equally weighted in the overall assessment? Some of the ten would seem to be more important than others (e.g., compare #7 to #6). This uncertainty problem is exacerbated for other corruption indicators for which no such criteria are made public at all, as is the

case for corruption measures produced by two competitors of the ICRG: the Economics Intelligence Unit (EIU) and World Markets Research Centre (WMRC).⁹

A final distinction among corruption indicators is that some are more suitable than others 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 have no 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.¹⁰ Nations in Transit provides only a generally-worded set of criteria for each of its 1-7 ratings levels, written to apply not only to corruption but to NIT's six other indicators.¹¹ 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 and change over time. In principle, the CPIA criteria are fixed and explicit, so can be used to assess progress over time. In practice, however, the criteria are revised somewhat every few years, and they are sufficiently subjective that the standards for a given ratings value may not be fixed.

⁹ Unlike the case with ICRG, the EIU and WMRC corruption indicators are not included in their standard products, but are part of a set of "customized" indicators available for a separate fee.

¹⁰ Moreover, there is a dramatic and unexplained "break" in the data between October and November 2001. In a typical month few ratings change; for example from July 2001 to August 2001 the only change was a decline for Mexico. From October 2001 to November 2001, however, there were 10 increases and 43 decreases.

¹¹ 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.

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.¹² 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.

3. Composite Corruption Indexes

There are at least three possible justifications for constructing a single corruption index from multiple, distinct sources of corruption indicators. The first motive 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.

¹² The mean number of responses per country increased from 84 in 2004 to 94 in 2005.

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, 2005). Although the statistical methods vary somewhat, both of these indexes standardize corruption indicators from numerous sources to place them on a comparable scale, and compute an average (unweighted for TI, weighted for WBI) of them 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 with data available from even one of the many sources used.¹³

The original purpose for the TI index was to raise awareness of corruption, and to provide researchers with better data for analyzing the causes and consequences of corruption. It has achieved these goals in spectacular fashion, and is regularly cited in news reports on corruption around the world. The WBI index, appearing several years after the TI index, was intended by its authors to improve and expand on TI in several ways. 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 on various grounds (Lambsdorff, 2005a). Third, using many of the same data sources, WBI constructs five other broad "governance" indexes.¹⁴ Fourth, WBI weights available sources differently, in contrast to the equal weighting in TI of available sources for each

¹³ The index on TI's web site lists only countries for which three or more data sources are available. The index on Johan Lambsdorff's web site lists index values for all countries with available data on one or more sources. Lambsdorff is the creator of the TI index. See <http://www.icgg.org/corruption.index.html>

¹⁴ These five include Rule of Law, Voice and Accountability, Political Stability and Violence, Regulatory Quality, and Government Effectiveness.

country.¹⁵ 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 (1) covered by more data sources, and (2) 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. 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

¹⁵ Even for TI, some sources will have a greater overall weight in the full set of index values, merely by having data available for more countries. For example, WEF covers many more countries than does IMD.

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.¹⁶

Conceptual imprecision, uncertainty and inconsistency

The TI and WBI indexes by construction are even more *conceptually imprecise* than some of their broadly-defined components (e.g. ICRG, NIT and CPIA). They are also more *conceptually uncertain*: the uncertainty in how criteria are weighted is compounded by aggregation. 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, but instead 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 values changes over time. Moreover, the sources used in constructing the indexes vary from country to country in a given year. Estonia’s 6.4 corruption rating and Latvia’s 4.2 corruption rating in the 2005 TI index are based on two different sets of indicators, hence on differing implicit definitions of corruption.

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

¹⁶ The CPIA indicator used in the WBI index will be available for the first time in the summer of 2006, but only for the IDA-eligible (i.e. lower income) countries.

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 11.¹⁷

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 simply compare two countries source by source, not bothering to construct a composite index at all. Either of these options requires access to the underlying data, however, which neither TI nor WBI provide.

The definitional inconsistency across countries entailed by using a different mix of sources is the price of maximizing the number of countries covered by the index. Corruption ratings are generated for more countries by TI and WBI when more sources (even those with spotty coverage) are aggregated, but for any pair of countries the index values are very likely to reflect differing implicit definitions of corruption.

Tracking Changes Over Time

¹⁷ This information was provided by Jim Anderson, who closely examined data sources in the TI and WBI indexes.

The standardization procedure used to place different indicators on a common scale precludes the ability to track changes meaningfully over time. 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 (1996, 1998, 2002 and 2004). Not only index values, but even rankings are not comparable across years as 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 of other countries, and not for year-to-year changes in the underlying data sources and indicators available for the country in question. Even setting aside expanded country coverage, changes from one year to the next in a country's ratings on TI or WBI could be driven purely by adding a new source to the index, or dropping an outdated one. No TI index value for any ECA country was based on the same set of sources in both 2004 and 2005.¹⁸ The WBI indexes for 2002 and 2004 are based on the same set of sources for only 4 of the 27 countries in

¹⁸ The BEEPS data from 2002 were included in the TI index for 2004, but no BEEPS data were used in the 2005 index. For most ECA countries, this was not the only change in TI sources used in 2004 and 2005.

the ECA region: Hungary, Slovak Republic, Slovenia, and Tajikistan. As with a pairwise 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 of them 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).

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.¹⁹ However, many of their sources clearly are not independent. 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

¹⁹ Setting aside the technical problems, Lambsdorff (2005a, 2005b) at least 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* which demonstrate that “countries can substantially improve” their quality of governance “even in the short term.”

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. Interdependence does not even require, however, that sources directly check each others' ratings; it could also result merely from sources relying on the same publications for their information about conditions in countries.

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, however, 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 deliberately are 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, so many of the same executives are likely to be included in both sets of surveys.²⁰

²⁰ The WEF and IMD have at least one partner institute in common in 12 of the 51 countries that are included in both sets of surveys. These two organizations jointly published the "World Competitiveness

It is impossible to determine quantitatively the degree of interdependence among sources used in TI and WBI. Many of the cross-country or over-time differences they classify as “statistically significant” undoubtedly would not be, if the appropriate corrections for interdependence could be made.

This unknown but substantial degree of interdependence among many of the sources also obviates any 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 comprised of a firm survey, a household survey and an expert assessment may provide a richer set of information than 9 sources, if all 9 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.²¹ 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. Claims of “being precise about imprecision” (Kaufmann, Kraay and Zoido-Lobaton, 2000)) depend on independence of assessments, hence cannot be supported.

The Choice of Weights in Aggregation

Report” from 1989 through 1995, but went their separate ways in 1996, with the WEF publishing the “Global Competitiveness Report” and the IMD publishing the “World Competitiveness Yearbook.” Both organizations list their partner institutes on their web sites.

²¹ The TI index uses the most recent three years of data for WEF and IMD. The WBI index uses only the most recent year.

Simplicity, objectivity, transparency and replicability all argue for weighting each variable (or each source, for sources that provide multiple indicators are provided by one source) equally 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 the WMRC.

The goal of accuracy could justify differential weighting, if there is good reason to believe that some sources are more informative than others. For this reason, the WBI index weights some sources more heavily than others. Specifically, the 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 (e.g. if a source’s assessment criteria include factors other than corruption).²² 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 their accuracy. In that case, any truly independent source will appear to be relatively inaccurate: using different information or a different

²² For example, the criteria for the corruption indicator from Business Environmental Risk Intelligence (BERI) include “xenophobia.” The WBI index includes it, but TI excludes it based on this extraneous content. The authors of the WBI index emphasize measurement error rather than conceptual mismatch as their justification for weighting more heavily the sources correlated more strongly with others.

methodology, it 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. 1/4 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.²³

The availability of the composite indexes themselves can aggravate these 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

²³ Any additional source receiving a positive weight inevitably must reduce the collective weight of previous sources. However, addition of a non-independent source will reduce disproportionately the weight given to an independent source, by reducing the latter’s average agreement with the other sources.

corruption ratings, but in most months very few ratings are changed. September to October 2001 was typical, with a single ½ point change for Switzerland. From October to November, however, 47 ratings were reduced and 10 increased. Such a dramatic reassessment had not occurred in the ICRG either before (dating back to 1984) or since. As shown in Figure 1, the month-to-month correlations in ICRG always exceed .99, but fell to .88 in November 2001. Although the ICRG has not responded to repeated requests for an explanation for this break in the data, there is some evidence that ratings were re-adjusted to conform much more closely to the TI rankings. As shown in Figure 1, the ICRG ratings were correlated with the TI 2001 ratings (released at end of June 2001) at only .72. However, the correlation with TI rose to .91 with the massive re-calibration by ICRG in November.²⁴ 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.²⁵ It also indirectly suggests a problem for TI (and WBI), to the extent that ICRG may not be unique among sources in sometimes free riding on the assessments of other sources - including the TI and WBI indexes themselves - rather than basing assessments on their own independent information.²⁶

Fortunately, correlation with other sources is not the only proxy for accuracy that could be used in assigning differential weights in index construction. Some more

²⁴ Over the last four years, this correlation has gradually declined and recently was at .79.

²⁵ This problem is relatively minor, as the WBI index uses many other sources in addition to the ICRG and those used by TI.

²⁶ Surowiecki (2004) makes an analogous argument that more information of the wrong kind can contribute to stock market bubbles and crashes. The advent of cable television and the Internet, including particularly CNBC, “magnified the dependent nature of the stock market because it bombarded investors with news about what other investors were thinking.” 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.

plausible weighting schemes for a broadly-defined composite index of corruption

include:

- Weight more heavily those sources that represent truly independent assessments. The BEEPS would thereby receive a greater weight than WEF or ICRG. Weighting each type of source equally, as suggested above, is consistent with this reasoning.
- 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 survey sources, 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.
- Weight indicators based on conceptual grounds; e.g. if an equal mix of administrative corruption and “state capture” is desired but most available indicators pertain to the former, weight more heavily those that pertain to the latter.

A disadvantage to most of these 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 of these weighting choices 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.

4. Levels and Trends in Corruption for ECA Countries

With the review of corruption indicators in section 3 as background, this section reports evidence from BEEPS and other sources on corruption trends between 2002 and 2005. We compare ECA to other regions, and compare ECA countries to each other.

Table 2 reports summary statistics for the corruption variables included in both the 2002

and 2005 BEEPS.²⁷ Figures reported represent means (or proportions, in the last two columns), weighting each of the 27 countries equally.²⁸

The most dramatic improvement between 2002 and 2005 is in the “bribe tax,” which fell by one-third from 1.6% of firm revenues to 1.1%. The bribe tax reported is skewed across firms, with a majority of firms reporting 0% in both years. A positive value for “bribe tax” was reported by 44% of firms in 2002, declining to 37% in 2005.²⁹

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% of firms on average in 2002, falling to about 18% in 2005. About 26% 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% 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.

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.

²⁷ Based on BEEPS 2005 data, Anderson and Gray (2006) report corruption levels for 2005, and changes between 2002 and 2005. Using BEEPS 2002 data, Gray, Hellman and Ryterman (2004) report corruption levels for 2002, and changes between 1999 and 2002. Fewer changes in BEEPS contents and methods occurred in the 2005 survey than in the 2002 survey, making over-time comparisons for 2002 and 2005 more reliable than comparisons between 1999 and 2002. For example, the “bribe tax” question was worded differently in 1999 and 2002, but identically in 2002 and 2005. Serbia/Montenegro and Tajikistan were added to the survey in 2002, so can be included in comparisons of BEEPS II and III, but not of BEEPS I and II.

²⁸ Only 26 countries are included in the “bribe tax” comparison, as there were problems with data for Turkey in 2005.

²⁹ The question asked “what percent of total annual sales do firms like yours typically pay,” which should elicit more candid responses than if it were phrased specifically in terms of the respondent’s firm.

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, the share of firms reporting a significant impact on their business from Parliamentarians receiving bribes to affect their votes is little changed.³⁰ A slight improvement is evident for a similar question on payments to Government officials to affect the content of government decrees.

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.

This overall progress also hides uneven progress across countries.³¹ Country-level data on four BEEPS measures are reported in Tables 3 and 4. Table 3 shows country-level means for bribe tax (Q40), bribe frequency (Q39a), costs of state capture (Q44b, on payments to government officials to influence rulemaking), and corruption as a serious obstacle to doing business (Q54q).³² Figures in parentheses show ranks, for 2002 and 2005 respectively, among the BEEPS countries. Table 4 shows the average change over time for the same four variables, and in parentheses shows ranks in terms of change, with “1” showing the largest improvement and “27” the largest deterioration.

³⁰ The question does not specify whether the “private payments/gifts” to Parliamentarians were paid by the firm or by other firms, but simply refers to the “impact on your business” from such practices. Survey questions are printed in full in Appendix B.

³¹ 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).

³² These four were selected from the larger set primarily for their prominent use in Gray, Hellman, and Ryterman (2004). Several of the “state capture” variables they used were not included in the 2005 BEEPS; among the remaining state capture items Q44b was selected because firms cited it more frequently than Q44a (on payments to Parliamentarians), and it is correlated most highly with an overall index of the three state capture measures in the 2005 BEEPS.

The multidimensionality of corruption is apparent from these tables, suggesting the difficulty in concluding that “corruption is worse” in country X than in country Y. Macedonia ranked 5th in “bribe tax” in 2005, but 27th on corruption as an obstacle to doing business. Latvia ranks 4th on bribe frequency and 20th on state capture. However, the various measures are significantly and positively correlated, and there are some countries that rank consistently high or low across measures. Slovenia ranks 1st on three measures and 3rd on the other measure. Estonia ranks 2nd on three measures, and 6th on the other one. Albania ranks 24th or worse on all four measures. Azerbaijan never ranks higher than 20th, and Kyrgyz never higher than 21st.

A second pattern concerns changes over time. Different measures often move in opposite directions for a given country. Georgia is among the few countries showing largest improvements on all four measures, while Azerbaijan was among those showing large deteriorations on all four variables. If changes were independent across the four variables, we would expect only about three countries to show either only positive, or only negative, changes on all four measures. In fact, 9 countries show decreasing corruption on all four measures, and 2 others show increasing corruption on all four. Other than Georgia, the most striking cases of improvements in corruption are for Slovak Republic, Romania and Bulgaria. Slovenia and Estonia also show impressive improvement, given that they already had relatively low levels of corruption in 2002, as measured by all four indicators. Azerbaijan and Lithuania exhibit increasing corruption on all four measures. Kyrgyz Republic’s deterioration on three of the measures must also be considered disappointing, despite a large improvement in the bribe tax from a region-worst 3.7% in 2002 to a second-worst 2.5% in 2005.

Trends in other data sources

Three key distinctions should be kept in mind in comparing trends from BEEPS to trends in other assessments of corruption, for 2002 to 2005. First, most other sources do not “unbundle” corruption across various functions of government, but provide only a single broadly-defined indicator. Second, broadly-defined indicators from other sources will differ from BEEPS in their inclusion of other aspects of corruption, in addition to corruption in firm-state interactions. 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 7 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.³³ in 2002 was 3.11, increasing

³³ Slovenia was a 28th ECA country in the CPIA until it “graduated” in 2003. It is therefore excluded from comparisons here.

to 3.30 in the 2004 ratings.³⁴ As shown in Table 6, 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 64th in 2002, improving to 61st 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 82nd to 76th.

The **World Economic Forum** data (WEF) included only 14 ECA countries among a total of 79 with available data for both 2002 and 2005. These 79 include many developed countries. Table 5 reports on nine WEF variables, all scaled from a low value of 1 to a high value of 7. Trends in these nine 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 45th in 2002 to 51st in 2005.

³⁴ 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.

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).

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 92nd to 88th.³⁵ Of course, evidence

³⁵ 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/>).

from the TI index is at least partly redundant, because it includes the WEF, NIT and EIU measures discussed above.

The WBI index is constructed only for even-numbered years, so comparisons were made only over the period 2002-2004. Even over this shorter period, the average ranking for the 28 ECA countries improved from 111th to 106th. 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. The WBI index includes WEF, NIT, the CPIA, the ICRG, the EIU and the BEEPS. However, both the 2002 and 2004 WBI indexes use the same 2002 data from BEEPS, in theory imparting a status quo bias in WBI ratings for ECA countries. In practice, however, any bias is trivial, as the weight assigned to the BEEPS data by the WBI methodology is extraordinarily small in both 2002 and 2004.³⁶

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 on which countries in the region experienced the most improvement. Table 7 shows very modest correlations in changes from 2002 to 2005 among the expert-based assessments; not all of them are even of the “correct” sign. A closer look at the data reveals that:

- In **ICRG**, the largest increases from 2002 to 2005 were for Russia and Serbia/Montenegro (both from 1 to 2 on the 0-6 scale). Czech Republic showed a decline from 3 to 2.5. Poland and Hungary were unchanged.
- In **NIT**, the only improvement of ½ point or greater were all in the Balkans: Bosnia, Bulgaria, Macedonia, and Romania. Poland and Belarus had the largest deterioration, of ¾ point. Czech Republic and Hungary improved by ¼ point.
- The **CPIA**, consistent with NIT, showed improvements in the Balkans, but also showed increases for Belarus and Tajikistan. Czech Republic increased by ½ point, while ratings were reduced for Hungary and Poland.

³⁶ The BEEPS weight is about one-sixth the weight given to WEF or ICRG, and less than 1/20th the weight given to Nations in Transit. In fact, the WBI corruption index is correlated with NIT at .96.

- The **EIU**, consistent with NIT and CPIA, shows deterioration for Poland. The four countries with improving ratings overlap little if at all, however, with the countries showing improvement in ICRG, NIT and CPIA.

These changes appear to be inconsistent, perhaps sufficiently so to cast doubt on the argument that expert-based assessments consult each other's ratings, or are otherwise based on very similar information. A partial answer to this puzzle is that ratings changes in expert-based systems 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.

"Regression to the mean" is commonplace among these corruption indicators. When ratings changes from 2002-2005 are regressed on initial (2002) levels, the coefficient on initial levels is negative and highly significant. This pattern holds for the ICRG and CPIA, and for most BEEPS and WEF measures. It does not hold for NIT or EIU; in those cases the coefficient on initial values is near 0 and does not approach statistical significance. The declines in NIT are concentrated in the former Soviet republics, which were already rated in 2002 as more corrupt than the European countries in ECA.

Regression-to-the-mean is exemplified by the cases of Czech Republic and Poland. 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.

Poland was ranked 2nd-best in NIT, tied for 2nd-best in CPIA, and tied for 1st in the EIU among ECA countries in 2002. It ranked in the middle of the pack in ICRG in both 2002 and 2005, and tends to rank just above the middle on most WEF and BEEPS measures. Poland's downgrading in NIT left it tied (with Slovakia) for 4th; its downgrading in CPIA left it in a tie for 6th (with three other countries), and its downgrading by EIU put it in the middle ranks, tied with three other countries, behind eight others and ahead of eight others.

If these examples of converging assessments are more the rule than the exception, one would expect inter-correlations among these expert sources to be higher in 2005 than in 2002. The data confirm this prediction: the mean of the six inter-correlations among ICRG, NIT, CPIA and EIU increases from .78 in 2002 to .85 in 2005.

A second prediction is that some of the weak correlations in changes among sources will strengthen, when we control for initial levels. As shown in Table 7, changes in EIU are completely uncorrelated with changes in ICRG and CPIA. However, in multivariate regressions, the EIU change is found to be significantly correlated with the ICRG change controlling for initial (2002) levels of EIU and ICRG. A similar result is obtained in another regression substituting CPIA for ICRG.

In summary, other sources tend to agree with evidence from the BEEPS that corruption in the region declined from 2002 to 2005. The various sources often disagree on which countries experienced improvement or deterioration. Part of this inconsistency is only apparent, however, as disagreement among sources on direction of *changes* often is necessary to achieve greater convergence in *levels*. Any convergence of this sort represents a likely reduction in measurement error. A more fundamental explanation for

apparent disagreement among corruption indicators – in levels or changes or time – is that they do not all purport to measure exactly the same concept. These differences are easily seen in the definitions in Appendix A. Most notably, perhaps, the CPIA attempts to measure not only corruption in the public sector, but also “transparency and accountability.”

5. 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 be exploited to identify which aspects of corruption are best captured by broader, perception-based measures - including the BEEPS question on corruption as an obstacle to doing business. Table 8 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.³⁷

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 reported in Table 8 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

³⁷ These correlations are limited to ECA countries. The larger sample of WEF countries could be used to compute correlations with most of these indicators (all but BEEPS “obstacle” and NIT).

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 of those four is strongly correlated with bribes for utility connections, exports and imports, and tax collection. Their correlations with a WEF measure of favoritism in decision making 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 the measure 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 on the one hand, and the four broad indicators on the other, the highest correlation is .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 less to state-enterprise interactions, and more to misappropriation of taxpayer funds by government officials. One of these is titled "diversion of public funds," and the other "public trust in the financial honesty of politicians." Diversion of funds is most strongly correlated with EIU (.66), among the four expert-based indicators. It is most weakly correlated (.42) with CPIA - despite the fact that of the four only the CPIA explicitly includes diversion of funds in its definition. Public trust in honesty of politicians is correlated at .44 with ICRG, but only at .04 for CPIA.

The dimensions of corruption that present the largest obstacles to doing business are of course 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 5th 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 (.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% 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 .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. Correlations of the

composite indexes with BEEPS and WEF, reported in the final two columns of Table 8, must be interpreted differently. Many of the firm survey variables are correlated by construction with the composite indexes; these correlations are shown in bold in Table 8.

The WEF administrative corruption variables – in the bottom five in Table 8 - 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 we find in Table 8. Correlations of TI with the three state capture variables in BEEPS range from .11 to .28. Correlations with TI exceed .40 for 7 of the 9 BEEPS administrative corruption measures, with the highest for tax collection (.66) and business licenses (.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, however. State capture measures in BEEPS are even more weakly correlated with WBI than with TI, despite the fact that some of them are components of WBI (unlike the case for TI). Although they are components, their weight is extraordinarily small, only 1/24th of the weight given to NIT in the 2004 WBI index.

Due to this huge weight for NIT, the correlation of WBI with NIT is .96; it is not surprising therefore that results in Table 8 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. Corruption in procurement - as reported in firm surveys - has little to do with rankings on TI and WBI, however.

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 8) yields two significant factors (i.e., with eigenvalues exceeding one) that together explaining 85% 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), to deal with fire and building inspections (Q41e), and to 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 8 is substituted for NIT.

Factor analyses based on the 9 WEF variables in Table 8 produce similar findings to those based on BEEPS variables. Two significant factors explain 95% of the variation. These factors are again clearly identifiable as state capture and administrative corruption, with the first four WEF variables listed in Table 8 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 B for criteria).³⁸

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 8. 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 8 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.

6. Summary and Conclusions

The BEEPS and other sources of corruption data indicate that corruption in ECA overall is declining. The sources appear to differ somewhat on the magnitude of this decline, with the BEEPS indicating a more favorable trend. Discrepancies in magnitudes do not necessarily indicate inaccuracy in one or more sources, however, because of differences in the (explicit or implicit) definition of corruption. The BEEPS and WEF each contains multiple items on administrative corruption and state capture, and both

³⁸ Regardless of the prevalence of bribery or other forms of corruption, it may be unrealistic to expect Nations in Transit – a product of Freedom House, funded in part by the U.S. government and conservative American philanthropists – to rate autocratic regimes highly on any indicator.

sources show much more evidence of improvement in the former than in the latter.

Unbundling administrative corruption, both sources exhibit reductions in bribe paying in utility connections, tax collection, and in importing and exporting, little change in the area of public procurement, and an increase in corruption in dealing with courts.

The various data sources disagree somewhat on which countries in the region have experienced the most progress. The BEEPS shows dramatic improvement in Georgia between 2002 and 2005, but no other source corroborates it. Expert assessments show more progress for other countries, particularly in the Balkans, and the WEF did not include Georgia in its 2002 surveys. There is some evidence of convergence in assessments of the various sources between 2002 and 2005, which can account in part for apparent inconsistencies across sources on which countries are experiencing improvements. To some extent, these discrepancies in changes – as with levels - may be explained by differences in how different sources define corruption. As Gray, Hellman and Ryterman (2000: 50) conclude from the BEEPS I and II data: “One cannot simply say that corruption is going up or down in individual countries, as we find a complex web of movements and mutations across different forms, features and dimensions of corruption. We need to be cautious and modest and to constantly recognize the full complexity of the measurement effort.”

Gray, Hellman and Ryterman (2000: 40) attribute part of the decline in corruption measured in BEEPS I and II 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. 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. More research is needed concerning the impact of optimism, recent economic performance, and recent corruption scandals on country-level corruption indicators, of the expert-assessment type as well as firm and household surveys.

More research is also needed inquiring 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 to be measuring primarily administrative corruption. Evidence in this paper is based solely on the ECA region, but this issue could be examined further using the full WEF sample.

The conceptual and methodological discussions, as well as empirical evidence reported here from BEEPS and other sources, 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, recognizing that their conceptual breadth makes them appropriate for some purposes but not for others. There should be more

examination of the criteria and methods of their underlying sources, to better understand what they are measuring, and to determine (roughly) 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. Also, the weights used by TI and WBI are essentially arbitrary, particularly when it is acknowledged that many of the data sources are highly interdependent. 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 in a large set of survey questions. Firm-level analyses, e.g. on firm characteristics associated with different forms of corruption, can also be conducted using BEEPS, although not with WEF. An important limitation of firm surveys such as BEEPS however is that it “provides a very incomplete measurement of corruption” (Gray, Hellman, and Ryterman, 2000: 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 margins. 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 PEFA indicators on public expenditure management.³⁹

³⁹ Information on these initiatives can be found, respectively, at <http://www.globalintegrity.org/>, <http://www.internationalbudget.org/>, and <http://www.pefa.org/>.

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) Economic Intelligence Unit (EIU) World Markets Research Centre (WMRC) World Bank CPIA
surveys of "well-informed persons" within country	UNECA African Governance Indicators World Governance Assessments
Composite indexes	
aggregation from various sources	TI corruption perceptions index WBI control of corruption index

Table 2
BEEPS 2002-2005 Summary statistics (weighting each country equally)

	Mean 2002	Mean 2005	% 2002	% 2005
	% of sales		% greater than 0	
Bribe tax (Q40)	1.59	1.05	43.6	36.8
Kickback for govt. contract (Q42)	1.99	1.85	27.4	25.6
	1-4 scale		% moderate or major obstacle	
Corruption problematic for business (Q54q)	2.24	2.14	21.0	17.8
	1-6 scale		% frequently, usually or always	
Bribe frequency (Q39a)	2.61	2.35	25.6	20.3
Bribe predictability (Q39b)	2.64	2.38	30.0	24.6
Utilities (Q41a)	1.58	1.47	6.3	4.5
Licenses & permits (Q41b)	2.08	2.01	15.9	13.9
Government contracts (Q41c)	1.93	1.96	15.9	16.4
Health/safety inspections (Q41d)	1.80	1.71	10.8	8.4
Fire/building inspections (Q41e)	1.87	1.75	11.4	9.0
Environmental inspections (Q41f)	1.67	1.56	7.9	6.0
Taxes & tax collections (Q41g)	2.08	1.96	17.5	14.3
Customs/imports (Q41h)	1.90	1.73	15.4	11.5
Courts (Q41i)	1.63	1.66	8.4	9.4
Influence legislation/rules (Q41j)	1.42	1.41	5.0	4.7
	1-5 scale		% moderate, major or decisive impact	
Impacted of capture: parliament (Q44a)	0.37	0.36	10.1	10.3
Impact of capture: govt. officials (Q44b)	0.44	0.38	12.4	11.6

Table 3
Corruption levels in ECA, from 2005 BEEPS

	Means or proportions* (2002, 2005 ranks in parentheses)			
	Bribe tax (Q40)	Bribe freq. (Q39a)	Impact of capture (Q44b)	Obstacle (Q54q)
Albania	1.80 (25, 24)	0.46 (23, 26)	0.88 (24, 27)	0.32 (27, 25)
Armenia	1.17 (7, 20)	0.10 (5, 7)	0.47 (1, 17)	0.15 (7, 10)
Azerbaijan	2.89 (24, 26)	0.27 (16, 22)	0.57 (10, 23)	0.21 (13, 20)
Belarus	1.11 (15, 19)	0.22 (15, 16)	0.04 (2, 1)	0.07 (12, 3)
Bosnia	0.39 (9, 3)	0.20 (11, 13)	0.84 (22, 26)	0.25 (24, 22)
Bulgaria	1.58 (17, 23)	0.16 (18, 12)	0.51 (27, 19)	0.19 (19, 16)
Croatia	0.76 (2, 10)	0.11 (3, 9)	0.26 (14, 10)	0.18 (16, 15)
Czech. Rep.	0.63 (6, 6)	0.10 (4, 6)	0.38 (7, 12)	0.20 (6, 19)
Estonia	0.29 (1, 2)	0.06 (2, 2)	0.17 (8, 6)	0.04 (1, 2)
Georgia	0.46 (23, 4)	0.07 (25, 3)	0.16 (21, 4)	0.20 (26, 17)
Hungary	1.06 (10, 15)	1.06 (10, 15)	0.19 (5, 8)	0.11 (3, 7)
Kazakhstan	1.42 (19, 21)	0.10 (12, 5)	0.19 (3, 7)	0.12 (9, 8)
Kyrgyz Rep.	2.46 (26, 25)	0.53 (27, 27)	0.45 (16, 21)	0.33 (15, 26)
Latvia	0.71 (8, 9)	0.07 (7, 4)	0.46 (23, 20)	0.10 (5, 5)
Lithuania	0.87 (3, 12)	0.24 (10, 20)	0.46 (11, 22)	0.14 (10, 9)
Macedonia	0.62 (4, 5)	0.25 (13, 21)	0.73 (26, 25)	0.35 (23, 27)
Moldova	1.09 (18, 18)	0.22 (19, 17)	0.45 (13, 15)	0.20 (20, 18)
Poland	0.70 (11, 8)	0.15 (8, 11)	0.33 (6, 13)	0.16 (18, 12)
Romania	0.81 (21, 11)	0.23 (24, 18)	0.36 (15, 14)	0.29 (25, 24)
Russia	1.07 (12, 16)	0.39 (26, 25)	0.29 (4, 11)	0.17 (8, 13)
Serbia/Mont.	0.67 (16, 7)	0.33 (6, 24)	0.61 (20, 24)	0.26 (11, 23)
Slovak Rep.	0.93 (13, 13)	0.11 (22, 8)	0.22 (17, 5)	0.11 (21, 6)
Slovenia	0.17 (5, 1)	0.05 (1, 1)	0.16 (12, 3)	0.04 (2, 1)
Tajikistan	1.07 (22, 17)	0.21 (21, 14)	0.48 (19, 18)	0.16 (14, 14)
Turkey	--	0.13 (14, 10)	0.39 (25, 16)	0.17 (17, 14)
Ukraine	1.52 (20, 22)	0.28 (20, 23)	0.20 (9, 9)	0.23 (22, 21)
Uzbekistan	0.99 (14, 14)	0.21 (9, 15)	0.10 (18, 2)	0.09 (4, 4)

*Proportions (as defined in Table 2) are reported for bribe frequency, state capture and corruption as an obstacle.

Table 4
Corruption trends in ECA, 2002-2005 BEEPS
(a – or + respectively indicates a reduction or increase in corruption)

	Changes in means or proportions (ranks in parentheses)			
	Bribe tax (Q40)	Bribe freq. (Q39a)	Impact of capture (Q44b)	Obstacle (Q54q)
Albania	-1.51 (4)	+0.20 (24)	+0.10 (27)	-0.16 (2)
Armenia	+0.25 (26)	-0.23 (16)	+0.10 (26)	+0.01 (20)
Azerbaijan	+0.15 (25)	+0.06 (23)	+0.07 (25)	+0.01 (21)
Belarus	-0.38 (16)	-0.29 (15)	-0.03 (9)	-0.11 (4)
Bosnia	-0.56 (11)	+0.01 (21)	+0.05 (24)	-0.10 (5)
Bulgaria	-0.37 (16)	-0.60 (5)	-0.16 (1)	-0.06 (8)
Croatia	+0.12 (23)	-0.10(18)	-0.05 (8)	-0.04 (13)
Czech. Rep.	-0.29 (18)	-0.12 (17)	+0.04 (21)	+0.08 (25)
Estonia	-0.05 (21)	-0.43 (8)	-0.02 (10)	-0.01 (18)
Georgia	-2.28 (1)	-1.42 (1)	-0.09 (5)	-0.15 (3)
Hungary	+0.10 (22)	-0.35 (9)	+0.01 (15)	+0.02 (22)
Kazakhstan	-0.68 (8)	-0.31 (11)	-0.00 (13)	-0.02 (15)
Kyrgyz Rep.	-1.24 (5)	+0.35 (26)	+0.03 (20)	+0.12 (27)
Latvia	-0.22 (19)	-0.52 (7)	-0.09 (3)	+0.02 (16)
Lithuania	+0.14 (24)	+0.33 (25)	+0.02 (17)	+0.02 (17)
Macedonia	-0.17 (6)	+0.01 (20)	+0.00 (14)	+0.03 (24)
Moldova	-0.99 (6)	-0.35 (10)	+0.01 (16)	+0.06 (10)
Poland	-0.51 (13)	-0.31 (12)	+0.05 (23)	-0.09 (6)
Romania	-1.75 (2)	-0.72 (4)	-0.02 (11)	-0.06 (9)
Russia	-0.36 (17)	+0.05 (22)	+0.04 (22)	+0.03 (23)
Serbia/Mont.	-0.85 (7)	+0.50 (27)	+0.03 (18)	+0.09 (26)
Slovak Rep.	-0.52 (12)	-0.77 (2)	-0.11 (2)	-0.17 (1)
Slovenia	-0.62 (10)	-0.30 (14)	-0.06 (7)	-0.02 (14)
Tajikistan	-1.52 (3)	-0.53 (6)	+0.03 (19)	-0.04 (12)
Turkey	--	-0.75 (3)	-0.09 (4)	-0.07 (7)
Ukraine	-0.67 (9)	-0.30 (13)	-0.01 (12)	-0.05 (11)
Uzbekistan	-0.46 (14)	+0.01 (19)	-0.08 (6)	-0.01 (19)

Table 5
Corruption Trends for ECA in Non-BEEPS Sources

Source	N ECA, all	Scale	Mean value		Mean rank	
			2002	2005	2002	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
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, 184	0 +/- sd	-0.41	-0.37	111	106

*NIT is the only source for which larger values indicate more corruption. Median value for NIT was 5.25 in 2002, improving to 5 in 2005. Average change in CPIA Q16, and in WBI control of corruption, are for 2002-2004.

Table 6
Changes in CPIA Q16, 2002-2004, by region

	Change 2002-4	2004	2002
EAP (18)	+0.32	3.12	2.79
ECA (27)	+0.19	3.30	3.11
AFR (45)	+0.12	2.95	2.83
LAC (28)	+0.09	3.54	3.45
SAR (7)	+0.00	3.21	3.21
MNA (9)	-0.06	2.89	2.95
All (134)	+0.14	3.18	3.04

Table 7: correlations in changes from 2002 to 2005

	NIT	ICRG	CPIA
ICRG	-.34	--	--
CPIA	-.34	<i>-.17</i>	--
EIU	+.31	<i>-.02</i>	<i>+.02</i>

Higher values indicate more corruption in the NIT and EIU scales, but less corruption on ICRG and CPIA. Correlations in the “wrong” direction are shown in italics.

Table 8
What Aspects of Corruption do Broad Perceptions Measures Capture?
Evidence from the 2005 BEEPS and WEF

	NIT	ICRG	CPIA	EIU	Obstacle	TI	WBI
BEEPS							
Bribe tax (Q40)	.54	-.50	-.44	.52	.37	-.48	-.51
Bribe frequency (Q39a)	.57	-.52	-.54	.80	.60	-.60	-.58
Bribe predictability (Q39b)	.56	-.48	-.52	.65	.52	-.61	-.58
Obstacle to business (Q54q)	.31	-.35	-.12	.50	--	-.46	-.33
Kickback in govt. contracts (Q42)	.14	-.04	-.04	.19	.33	-.15	-.12
Administrative corruption							
Utilities (Q41a)	.50	-.34	-.52	.60	.54	-.60	-.55
Licenses & permits (Q41b)	.68	-.58	-.66	.82	.59	-.70	-.70
Government contracts (Q41c)	-.08	-.07	<i>.21</i>	.01	.57	-.06	<i>.05</i>
Health/safety inspections (Q41d)	.05	-.33	<i>.02</i>	.38	.54	-.17	-.09
Fire/building inspections (Q41e)	.50	-.52	-.43	.58	.18	-.45	-.50
Environmental inspections (Q41f)	.43	-.42	-.34	.52	.53	-.40	-.41
Taxes & tax collections (Q41g)	.65	-.58	-.68	.79	.49	-.66	-.69
Customs/imports (Q41h)	.38	-.41	-.35	.58	.72	-.46	-.42
Courts (Q41i)	.30	-.40	-.22	.47	.67	-.41	-.33
State capture							
Influence legislation/rules (Q41j)	.03	-.20	<i>.03</i>	.09	.62	-.11	-.04
Impact of capture: parliament (Q44a)	.15	-.39	-.08	.36	.68	-.28	-.19
Impacted of capture: govt. off. (Q44b)	.13	-.29	-.06	.32	.73	-.27	-.17
WEF							
Favoritism in decisions	-.34	.49	.28	-.43	-.52	.52	.39
Diversion of public funds	-.52	.56	.42	-.66	-.83	.67	.58
Business costs of corruption	-.72	.65	.72	-.79	-.68	.80	.79
Financial honesty of politicians	-.22	.44	.04	-.36	-.63	.40	.24
Irregular payments in...							
...exports & imports	-.73	.59	.77	-.68	-.57	.85	.79
...public utilities	-.73	.58	.78	-.73	-.55	.86	.81
...tax collection	-.74	.69	.77	-.68	-.49	.82	.80
...public contracts	-.26	.05	.27	-.10	-.40	.53	.35
...judicial decisions	-.69	.57	.67	-.57	-.55	.82	.75

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.

References

- James Anderson and Cheryl Gray (2006). *Anticorruption in Transition 3: Who is Succeeding, and Why?* Washington DC: The World Bank.
- Economic Commission for Africa (2005). *Striving for Good Governance in Africa*. Addis Ababa: UNECA.
- Gray, Cheryl; Joel Hellman and Randi Ryterman (2004). *Anticorruption in Transition 2: Corruption in Enterprise-State Interactions in Europe and Central Asia 1999-2002*. Washington DC: The World Bank.
- Hyden, Goran; Julius Court and Kenneth Mease (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, Daniel (2005). "Myths and Realities of Governance and Corruption." In A. Lopez- Claros, M. E. Porter and K., *Global Competitiveness Report 2005-2006: Policies Underpinning Rising Prosperity..* Houndmills, UK: Palgrave Macmillan for the World Economic Forum.
- Kaufmann, Daniel; Aart Kraay and Massimi Mastruzzi (2005). "Governance Matters IV: Governance Indicators for 1996-2004." World Bank Policy Research Working Paper 3630.
- Kaufmann, Daniel; Aart Kraay and Massimi Mastruzzi (2003). "Governance Matters III: Governance Indicators for 1996-2002." World Bank Policy Research Working Paper 3106.
- Kaufmann, Daniel; Aart Kraay and Pablo Zoido-Lobaton (2000). "Governance Matters: From Measurement to Action." *Finance and Development*, 37(2): 10-13.
- Lambsdorff, Johann Graf (undated). "Measuring the Dark Side of Human Nature: The Birth of the Corruption Perceptions Index" (http://www.icgg.org/corruption.cpi_childhooddays.html).
- Lambsdorff, Johann Graf (2005a). "The Methodology of the 2005 Corruption Perceptions Index" (http://www.icgg.org/downloads/CPI_Methodology.pdf).
- Lambsdorff, Johann Graf (2005b). "Determining Trends for Perceived Levels of Corruption." University of Passau Discussion Paper V-38-05.

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

Mauro, Paolo (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.

Sandholtz, Wayne and William Koetzle (2000). "Accounting for Corruption: Economic Structure, Democracy and Trade." *International Studies Quarterly*, 44, 31-50.

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

Swamy, Anand; Stephen Knack, Young Lee and Omar Azfar (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 A
Availability of data for both 2002 and 2005

	BEEPS	NIT	CPIA	WEF	ICRG
Albania	Y	Y	Y	2005	Y
Armenia	Y	Y	Y	2005	Y
Azerbaijan	Y	Y	Y	2005	Y
Belarus	Y	Y	Y	N	Y
Bosnia	Y	Y	Y	2005	N
Bulgaria	Y	Y	Y	Y	Y
Croatia	Y	Y	Y	Y	Y
Czech. Rep.	Y	Y	Y	Y	Y
Estonia	Y	Y	Y	Y	Y
Georgia	Y	Y	Y	2005	N
Hungary	Y	Y	Y	Y	Y
Kazakhstan	Y	Y	Y	2005	Y
Kyrgyz Rep.	Y	Y	Y	2005	N
Latvia	Y	Y	Y	Y	Y
Lithuania	Y	Y	Y	Y	Y
Macedonia	Y	Y	Y	2005	N
Moldova	Y	Y	Y	2005	Y
Poland	Y	Y	Y	Y	Y
Romania	Y	Y	Y	Y	Y
Russia	Y	Y	Y	Y	Y
Serbia/Mont.	Y	Y	Y	2005	Y
Slovak Rep.	Y	Y	Y	Y	Y
Slovenia	Y	Y	2002	Y	Y
Tajikistan	Y	Y	Y	2005	N
Turkey	Y	N	Y	Y	Y
Turkmenistan	N	Y	Y	N	N
Ukraine	Y	Y	Y	Y	Y
Uzbekistan	Y	Y	Y	N	N

Y = 2002 and 2005 available, N = neither year available

Appendix B: Definitions of Corruption Indicators

1. BEEPS Questions

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 percent 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 percent 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 U. S. 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 .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 government 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 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's 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.

- 1
 - a. 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.
 - b. Government decision-making is secretive. The public is prevented from participating in or learning about decisions and their implications.
 - c. The state has been captured by narrow interests (economic, political, ethnic, and/or military). Administrative corruption is rampant.
- 2
 - a. 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.
 - b. 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.
 - c. 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.
- 3
 - a. External accountability mechanisms such as inspector-general, ombudsman, or independent audit may exist, but have inadequate resources or authority.
 - b. 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.
 - c. Elected and other public officials often have private interests that conflict with their professional duties.
- 4
 - a. 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.
 - b. 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.
 - c. 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.
- 5
 - a. 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.
 - b. 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.
 - c. 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)

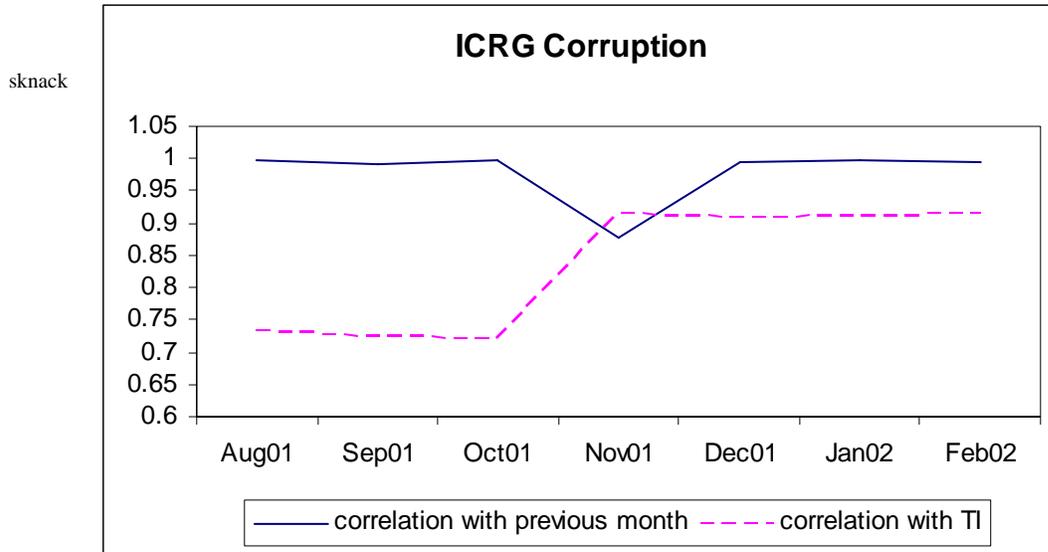
Diversion of public funds

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)

Figure 1



L:\BEEPS\wps_BEEPS_Apr25.doc
05/03/2006 9:51:00 AM