

Background Paper to the 2004 Corruption Perceptions Index

Framework Document 2004

The Corruption Perceptions Index (CPI) is a composite index, using data compiled between 2002 and 2004. 18 surveys of businesspeople and assessments by country analysts from 12 independent institutions enter the CPI.

All sources employ a homogeneous definition of “extent of corruption”. The assessments are gathered from experienced respondents and enhance our understanding of real levels of corruption.

Comparisons to last year’s index should be based on scores. However, such comparisons can be misleading because of methodological changes between years.

Non-parametric statistics are used for standardizing the data and for determining the precision of the scores.

Prof. Dr. Johann Graf Lambsdorff

Transparency International (TI) and
University of Passau
September 2004

1. The methodology

The year 2004 marks the tenth publication of the Transparency International Corruption Perceptions Index (CPI). Since 1995, the first publication of the CPI, there has been a wave of scientific publications based on the CPI and our knowledge is still expanding quickly. As in previous years, this framework document provides an in-depth explanation of the methodology and measurement precision.

The goal of the CPI is to provide data on extensive perceptions of corruption within countries. The CPI is a composite index, making use of surveys of businesspeople and assessments by country analysts. It consists of credible sources using diverse sampling frames and different methodologies. These perceptions enhance our understanding of real levels of corruption from one country to another.

As pointed out in previous framework documents, unbiased, hard data continue to be difficult to obtain and usually raise problematic questions with respect to validity. Comparing the number of prosecutions, for example, does not reflect actual levels of corruption but the quality of prosecutors. International surveys on perceptions therefore serve as the most credible means of compiling a ranking of nations.

Overall, 18 sources could be included in the 2004 CPI, originating from 12 independent institutions. The complete list of sources is presented in the appendix. All in all, the number of countries in the CPI increased from 133 to 146.

Sources in 2004

Prior to selecting sources guidelines have been set up which organize the underlying decision making process. These include the actual criteria that a source needs to meet in order to qualify for inclusion as well as organizational guidelines on how the final decision is reached with the help of the Transparency International Steering Committee.

This process aims at making the final decision as transparent and robust as possible. As a result of this it was decided that the 2004 CPI includes data from the following sources:

- **BEEPS**, The Business Environment and Enterprise Performance Survey, 2002.
- **CU**, the State Capacity Survey by the Center for International Earth Science Information Network (CIESIN) at Columbia University, 2003.
- **EIU**, The Economist Intelligence Unit, 2004.
- **FH**, Freedom House Nations in Transit, 2004.
- **II**, Information International, Beirut, Lebanon, 2003.
- **IMD**, The International Institute for Management Development, Lausanne. We will use the three annual publications from 2002-2004.
- **MDB**, a multilateral development bank, 2002.
- **MIG**, Grey Area Dynamics Ratings by the Merchant International Group, 2004.
- **PERC**, The Political and Economic Risk Consultancy, Hong Kong. We will use the three annual publications from 2002-2004.
- **TI/GI**, Gallup International on behalf of Transparency International, Bribe Payers Index Survey, 2002.
- **WEF**, The World Economic Forum. We will use the three annual publications from 2002-2004.
- **WMRC**, The World Markets Research Centre, 2004.

An essential condition for inclusion is that a source must provide a ranking of nations. This condition is not met if a source conducts surveys in a variety of countries but with varying methodologies. Comparison from one country to another would not be feasible in this case. Another condition is that sources must measure the overall extent of corruption. This is violated if aspects of corruption are mixed with issues other than

corruption such as political instability or nationalism or if changes are measured instead of the extent of corruption.

For example, the index "Corruption in Government" from the International Country Risk Guide (ICRG), conducted by the Political Risk Services (PRS), did not meet these requirements, albeit being widely used in research as a measure of the extent of corruption. It does not determine a country's level of corruption but the political risk involved in corruption. As pointed out to us by the ICRG-editor, these two issues can differ considerably, depending on whether there exists a high or low public tolerance towards corruption. Corruption only leads to political instability if it is not tolerated. Due to this, the data by PRS-ICRG did not qualify for inclusion in the CPI. However, TI hopes to include a modified set of data by PRS in the future.

The 2004 CPI combines assessments from the past three years to reduce abrupt variations in scoring that might arise due to random effects. Some sources, such as BEEPS, II, MDB and TI/GI provided only one recent survey. Others such as WEF, IMD and PERC conducted annual surveys between 2002 and 2004, which are all included.

While this averaging is valuable for the inclusion of surveys, it is inappropriate for application to the data compiled by professional risk agencies. Such assessments as compiled by EIU, CU, FH, MIG and WMRC are conducted by a small number of country experts who regularly analyze a country's performance, counterchecking their conclusions with peer discussions. Following this systematic evaluation, they then consider a potential upgrading or downgrading. As a result, a country's score changes rather seldom and the data shows little year-to-year variation. Changing scores in this case are the result of a considered judgment by the organization in question. To then go back and average the assessments over a period of time would be inappropriate. On the other hand, in the case of surveys of elite businesspeople an

averaging over various years produces a useful smoothing effect. While some annual data may contain random errors, these do not necessarily carry over to the next year.

Year-to-year comparisons

Comparisons to the results from previous years should be based on a country's score, not its rank. A country's rank can change simple because new countries enter the index and others drop out. A higher score is an indicator that respondents provided better ratings, while a lower score suggests that respondents revised their perception downwards. However, year-to-year comparisons of a country's score do not only result from a changing perception of a country's performance but also from a changing sample and methodology. Old sources drop out of the index and new sources enter, disturbing the consistency of the assessment. The index primarily provides an annual snapshot of the views of businesspeople and country analysts, with less of a focus on year-to-year trends.

However, to the extent that changes can be traced to a change in the assessments provided by individual sources, trends can be identified. Comparing older data (that is, data that was used in 2003¹ but no longer used this year) with topical data from 2004 allows us to identify such changes in perceptions during the last three years. Countries whose CPI score increased relative to 2003 and where this improvement is not the result of technical factors are Bahrain, Belize, Cyprus, Dominican Republic, Jamaica, Kuwait, Luxembourg, Mauritius, Oman, Poland, Saudi Arabia, Senegal and Trinidad and Tobago. The considerable decline in their scores does not result from technical factors - actual changes in perceptions are therefore likely.

With the same caveats applied, on the basis of data from sources that have been consistently used for the index, improvements can be observed for Austria,

¹ These data are EIU 2003, IMD 2001, WEF 2001, PERC 2001, FH 2003 and WMRC 2002.

Botswana, Czech Republic, El Salvador, France, Gambia, Germany, Jordan, Switzerland, Tanzania, Thailand, Uganda, United Arab Emirates and Uruguay.

2. Validity

All sources generally apply a definition of corruption such as the misuse of public power for private benefit, for example bribing of public officials, kickbacks in public procurement, or embezzlement of public funds. Each of the sources also assesses the “extent” of corruption among public officials and politicians in the countries in question:

- BEEPS asks senior businesspeople “thinking about officials ... It is common for firms in my line of business to have to pay some irregular “additional payments” to get things done”. (Always, Mostly, Frequently, Sometimes, Seldom, Never, Don’t know)” and “Using this scale (No Obstacle=1 ; Minor Obstacle=2 ; Moderate Obstacle=3 ; Major Obstacle=4 ; Don’t know/no answer=5) can you tell me how problematic are these different factors for the operation and growth of your business: ... corruption...”
- CU asks its panel of experts to rate the severity of overall corruption within the state on the following scale: Low; Low/Modest; Modest; Modest/Severe; Severe.
- EIU asks its panel of expert to assess the incidence of corruption and defines corruption as the misuse of public office for personal (or party political) financial gain. Integers between 0 (denoting a “very low” incidence of corruption) and 4 (denoting a “very high” incidence) are provided.
- FH asks its panel of expert to assess the implementation of anticorruption initiatives; the government’s freedom from excessive bureaucratic regulations and other controls that increase opportunities for corruption; public perceptions of corruption; the business interests of top policy makers; laws on financial disclosure and conflict of interest; audit and investigative rules for executive and legislative bodies; protections for whistleblowers, anticorruption activists, and others who report corruption; and the media’s coverage of corruption.
- II adopted a question similar to the one used by TI/GI.
- IMD surveys elite businesspeople and asks them to assess whether “bribing and corruption prevail or do not prevail in the economy.”
- MDB asks its staff to a foreign country with respect to the following questions: “How widespread is the incidence of corruption? (Widespread; Somewhat widespread; Somewhat limited; Limited; No judgment). The question is repeated for further foreign countries, depending on the respondents experience.
- MIG asks its panel of correspondents assess levels of corruption. Corruption in their definition ranges from bribery of government ministers to inducements payable to the “humblest clerk”.
- PERC asks expatriate businessmen to rate on a scale of zero to 10 how bad they considered the problem of corruption to be in the country in which they are working as well as in their home country.
- TI/GI asks “Which are the countries, besides this one, with which you have had the most business experience in the last 3-5 years? Please name up to five countries.
 - a. In [country 1], how common are payments (e.g. bribes) to obtain or retain business or other improper advantages to senior public officials, like politicians, senior civil servants, and judges?
 - b. In [country 1], how significant of an obstacle are the costs associated with such payments for doing business?
 - c. In [country 1], how frequently are public contracts awarded to business associates, friends and relatives rather than on a competitive bidding basis?”

Continue with countries 2-5. Scale for answers is from 'Very Common' [01] to 'Very Uncommon / Never'[04]. Don't know [88].

- WEF asks: "In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with:"
 - 1 – exports and imports
Common |1|2|3|4|5|6|7| Never occur
 - 2 - public utilities (e.g. telephone or electricity)
Common |1|2|3|4|5|6|7| Never occur
 - 3 - annual tax payments
Common |1|2|3|4|5|6|7| Never occur
 - 4 – public contracts
Common |1|2|3|4|5|6|7| Never occur
 - 5 - loan applications
Common |1|2|3|4|5|6|7| Never occur
 - 6 - influencing laws and policies, regulations, or decrees to favor selected business interests?
Common |1|2|3|4|5|6|7| Never occur
 - 7 – getting favorable judicial decisions
Common |1|2|3|4|5|6|7| Never occurFrom these questions the simple average has been determined.
- WMRC provides an assessment of the likelihood of encountering corrupt officials. Corruption can range from petty bureaucratic corruption (such as the paying of bribes to low-level officials) right through to grand political corruption (such as the paying of large kick-backs in return for the awarding of contracts). Scores take the following values: 1; 1.5; 2; 2.5; 3; 3.5; 4; 4.5; 5. They have the following meaning:
 1. This country will have an excellent business environment and corruption will be virtually unknown.
 2. This country will have a good and transparent business environment. Corruption - official and otherwise - may occur occasionally, but most businesses will not encounter this.
 3. This country will have some significant operational obstacles, including corruption. However, whilst official corruption may be relatively common, it should

not affect business in an overly negative manner.

4. This country will have a poor business environment. Corruption is likely to be endemic in the business world and officialdom, and it will not be uncommon for kick-backs or bribes to be demanded in return for the awarding of contracts.
5. This country will have severe operational obstacles, which in practice make business impossible. Corruption will be pervasive and will reach the highest levels of government.

The various terms used by the sources "prevalence", "commonness", "frequency", "likelihood", "problematic" and "severity" are closely related. They all refer to some kind of extent of corruption, which is also aim of the CPI. This common feature of the various sources is particularly important in view of the fact that corruption comes in different forms. It has been suggested in numerous publications that distinctions should be made between these forms of corruption, e.g. between nepotism and corruption in the form of monetary transfers. Yet, none of the data included in the CPI emphasize one form of corruption at the expense of other forms. The sources can be said to aim at measuring the same broad phenomenon. As also emphasized in the framework documents of previous years, the sources do not distinguish between administrative and political corruption, nor between petty and grand corruption.

The term "extent of corruption" may imply different things. In particular, it may relate to the frequency of bribes or the size of bribes. But we know from the results of our sources that frequency and the size of bribes tend to correlate highly (as expounded in the framework documents of previous years). In countries where corruption is frequent it also amounts to a high fraction of firms' revenues. In sum, the term "extent of corruption" seems to equally reflect the two aspects, frequency of corruption and the total value of bribes paid.

3. Samples, perceptions and reality

While the sources all aim at measuring the extent of corruption, the sample design differs considerably. Basically, three different types of samples are used.

A first group of sources, namely CU, EIU, FH, MIG and WMRC, assemble the perceptions of non-residents, turning in their experienced perception with regard to foreign countries. These assessments are largely carried out by respondents from developed countries of the western hemisphere such as North America and Western Europe.

A second group of sources, namely TI/GI, II and MDB, assembles also the perceptions of non-residents, but these respondents are largely from less developed countries. There is an advantage to perceptions vis-à-vis foreign countries because they are not vulnerable to a “home-country bias”. Such a type of bias would be relevant if respondents assess their home country purely according to local standards. Such a standard would be problematic because it can differ from one country to another, impairing the validity of cross-country comparisons.

A third group of sources, namely BEEPS, IMD, PERC and WEF, gather assessments made by residents with respect to the performance of their home country. These respondents are partly nationals but also expatriates from multinational firms. While such data might be susceptible to the aforementioned “home-country bias”, they are not susceptible to introducing an undue dominance of “western business people’s” viewpoint. Such a viewpoint would be inadequate if foreigners lack a proper understanding of a country's culture.

The data correlate well with each other, irrespective of these different methodologies. The high correlations ameliorate fears that any of the aforementioned biases are important to the results. Residents may therefore have a rather universal ethical standard and adequately position their country as compared to foreign countries.

Likewise, those respondents who assess foreign countries seem to have a good grasp of a country’s culture and provide appropriate assessments in the light of this.

The second group, consisting of II, MDB and TI/GI, is less susceptible to both biases. These sources either survey their local staff members (in the case of MDB) or respondents from emerging economies and less developed countries (TI/GI and II). In the latter case respondents are asked to assess the performance of industrial countries and neighboring countries. Those polled are not asked to assess their home country or individual foreign countries, but to provide a comparative assessment of various foreign countries. This approach makes sure that, first, a consistent ethical standard is applied to all countries, that, second, only those countries are assessed where sufficient experience and cultural insights are available and that, third, the viewpoints of respondents from less developed countries are well represented. Yet, as shown in the correlations, this different approach does not bring about noteworthy different results.

In sum, the perceptions gathered are a helpful contribution to the understanding of real levels of corruption. As was also explained in detail in the 2001 framework document, the perceptions gathered well relate to actual experience made and less to hearsay.²

4. The index

Standardizing

Each of the sources uses its own scaling system, requiring that the data be standardized before each country’s mean value can be determined. This standardization is carried out in two steps.

For step 1 each source is standardized using matching percentiles. The *ranks* (and not the scores) of countries is the only in-

² See Lambsdorff, J. Graf (2001) “Framework Document.”, Background Paper 2001 Corruption Perceptions Index:

ICGG.org/downloads/2001_CPI_FD.pdf

formation processed from each source. For this technique the common sub-samples of a new source and the previous year's CPI are determined. Then, the largest value in the CPI is taken as the standardized value for the country ranked best by the new source. The second largest value is given to the country ranked second best, etc.³ Imagine that a new sources ranks only five countries: UK (4.2), Singapore (3.9), China (2.8), Malaysia (2.7) and India (2.4). In the 2003 CPI these countries obtained the scores 8.7, 9.4, 3.4, 5.2 and 2.8, respectively. Matching percentiles would now assign UK the best score of 9.4, Singapore second best with 8.7, China 5.2, Malaysia 3.4 and India 2.8.

Matching percentiles is superior in combining indices that have different distributions. Not the cardinal information is processed but only the ordinal information provided by a source. Many of the alternative parametric standardization methods, on the other hand, would require a multitude of assumptions – some of which may not be realistic. But, as matching percentiles makes use of the ranks and not the scores of sources, this method loses some of the information inherent in the sources. What tips the balance in favor of this techniques is its

³ If two countries share the same rank, their standardized value is the simple mean of the two respective scores in the CPI. The scores for countries where no CPI value was available are determined by referring to neighbor countries in the source's ranking. Linear interpolation is applied to their scores, suggesting that if a source assigns such a country a score close to the upper neighbor, also its standardized value is closer to that of this neighbor. If such a country is ranked best (or worst) by a source it would have only one neighbor, not two. The second neighbor is constructed by using the highest (or lowest) attainable score by the source and the CPI value 10 (or 0). This approach guarantees that all values remain within the range between 10 and 0.

capacity to keep all reported values within the bounds from 0 to 10. This results because any standardized value is taken from the previous year's CPI, which by definition is restricted to the aforementioned range. This characteristic is not obtained by an alternative technique, e.g. one that standardizes the mean and standard deviation of the joint sub-samples of countries.

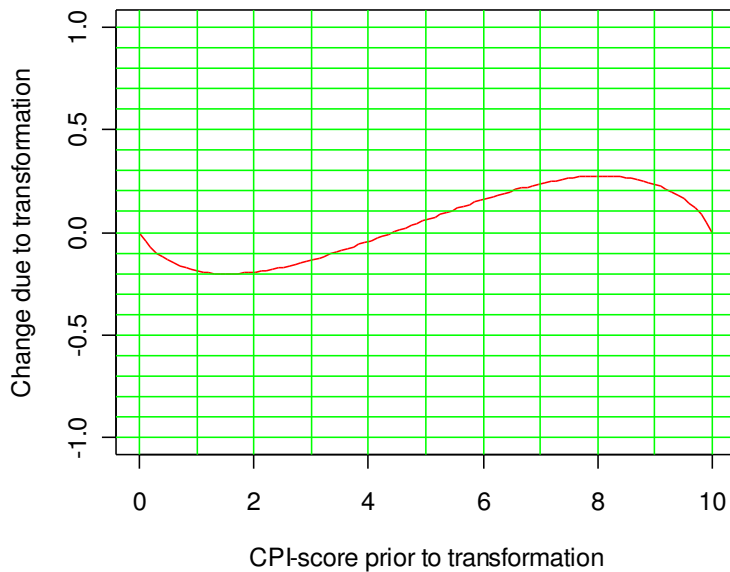
Step 2

Having obtained standardized values that are all within the reported range, a simple average from these standardized values can be determined. However, the resulting index has a standard deviation that is smaller than that of the CPI of previous years. Without a second adjustment there would be a trend towards a continuously smaller diversity of scores. If, e.g., Finland were to repeat its score from the previous year, it would have to score best in all sources. If it scores second to best in any source, the standardized value it obtains after using matching percentiles and aggregation would be lower than its current score. Thus, given some heterogeneity among sources, it seems inevitable that Finland's score would deteriorate. The opposite would be true of Bangladesh, which would obtain a better score if it is not consistently rated worst by all its sources. A second standardization is required in order to avoid a continuous trend to less diversity among scores.

However, simply stretching the scores (by applying a simple mean and standard deviation technique) might bring about values that are beyond our range from 0 to 10. A more complicated standardization is required for the second step: A beta-transformation. The idea behind this monotonous transformation is to increase the standard deviation to the previous year's value, while preserving the range from 0 to 10. Each value (X) is therefore transformed according to the following function:

$$10 * \int_0^1 (X/10)^{\alpha-1} (1 - X/10)^{\beta-1} dX$$

Figure 1: Beta Transformation



This beta-transformation is available in standard statistics programs. The crucial task is to find the parameters α and β so that the resulting mean and standard deviation of the index have the desired values. An algorithm has been determined that carries out this task. Applying this approach to the 2004 CPI, the change in the scores is depicted by figure 1. The parameters are $\alpha=1.168$ and $\beta=1.187$. As shown in the figure, scores between 4.5 and 10 are increased slightly, while those between 0 and 4.5 are lowered.

The beta transformation is first applied to all values that were standardized in step 1. Afterwards the average of these are computed to determine a country's score. In our publication we also report the high-low range. This refers to all standardized values after carrying out the beta-transformation. This procedure ensured that the high-low range is consistently related to a country's mean value.

Reliability and Precision

A ranking of countries may easily be misunderstood as measuring the performance of a country with absolute precision. This is certainly not true. Since its start in 1995 TI has provided data on the standard deviation

and the amount of sources contributing to the index. This data already serves to illustrate the inherent imprecision. Also the high-low range is provided in the main table. This depicts the highest and the lowest values provided by our sources, so as to portray the whole range of assessments. However, no quick conclusions should be derived from this range to the underlying precision with which countries are measured. Countries which were assessed by 3 or 12 sources can have the same minimum and maximum values, but in the latter case we can feel much more confident about the country's score. In order to arrive at such measures of precision, other statistical methods are required.

The strength of the CPI is based on the concept that a combination of data sources combined into a single index increases the reliability of each individual figure. As in previous years, the 2004 CPI includes all countries for which at least three sources had been available. The idea of combining data is that the nonperformance of one source can be balanced out by the inclusion of at least two other sources. This way, the probability of misrepresenting a country is seriously lowered. This is valid even in case the sources are not totally independent of each other. Such partial dependency may arise if some respondents are aware of other people's perception of the level of corruption, or of other sources contributing to the CPI.

An indicator for the overall reliability of the 2004 CPI can be drawn from the high correlation between the sources. This can be depicted from the standard Pearson correlation and Kendall's rank correlation, provided in tables 1 and 2.

Table 1:¹⁾
Pearson
Correlation

	IMD 2003	IMD 2004	IMD 2002	PERC2002	PERC2003	PERC2004	GCR 2002	GCR 2003	GCR 2004	EIU 2004	FH 2004	TI/GI 2002	CU 2003	WMRC 2004	MIG	BEEPS02	Inf_Int2003	MDB 2002
IMD 2003	1.00	0.98	0.98	0.95	0.95	0.94	0.95	0.95	0.95	0.90	0.66	0.79	0.86	0.91	0.83	0.75	0.80	0.84
IMD 2004	0.98	1.00	0.98	0.92	0.94	0.94	0.93	0.95	0.95	0.91	0.58	0.81	0.86	0.91	0.87	0.74	0.74	0.87
IMD 2002	0.98	0.98	1.00	0.95	0.97	0.97	0.96	0.95	0.95	0.91	0.46	0.83	0.87	0.91	0.85	0.69	0.72	0.81
PERC2002	0.95	0.92	0.95	1.00	0.98	0.94	0.89	0.85	0.88	0.86		0.73	0.93	0.93	0.73		0.64	0.69
PERC2003	0.95	0.94	0.97	0.98	1.00	0.97	0.91	0.89	0.89	0.92		0.72	0.88	0.95	0.78		0.66	0.73
PERC2004	0.94	0.94	0.97	0.94	0.97	1.00	0.94	0.93	0.93	0.90		0.80	0.85	0.94	0.83		0.56	0.80
GCR 2002	0.95	0.93	0.96	0.89	0.91	0.94	1.00	0.95	0.94	0.88	0.37	0.84	0.83	0.90	0.84	0.49	0.74	0.67
GCR 2003	0.95	0.95	0.95	0.85	0.89	0.93	0.95	1.00	0.95	0.89	0.77	0.80	0.79	0.90	0.84	0.81	0.83	0.71
GCR 2004	0.95	0.95	0.95	0.88	0.89	0.93	0.94	0.95	1.00	0.88	0.74	0.79	0.85	0.91	0.82	0.73	0.78	0.68
EIU 2004	0.90	0.91	0.91	0.86	0.92	0.90	0.88	0.89	0.88	1.00	0.86	0.77	0.86	0.91	0.82	0.72	0.74	0.64
FH 2004	0.66	0.58	0.46				0.37	0.77	0.74	0.86	1.00		0.76	0.82	0.79	0.60		0.77
TI/GI 2002	0.79	0.81	0.83	0.73	0.72	0.80	0.84	0.80	0.79	0.77		1.00	0.82	0.81	0.76		0.40	
CU 2003	0.86	0.86	0.87	0.93	0.88	0.85	0.83	0.79	0.85	0.86	0.76	0.82	1.00	0.80	0.78	0.72	0.70	0.75
WMRC 2004	0.91	0.91	0.91	0.93	0.95	0.94	0.90	0.90	0.91	0.91	0.82	0.81	0.80	1.00	0.82	0.43	0.81	0.62
MIG	0.83	0.87	0.85	0.73	0.78	0.83	0.84	0.84	0.82	0.82	0.79	0.76	0.78	0.82	1.00	0.50	0.72	0.70
BEEPS02	0.75	0.74	0.69				0.49	0.81	0.73	0.72	0.60		0.72	0.43	0.50	1.00		0.08
Inf_Int2003	0.80	0.74	0.72	0.64	0.66	0.56	0.74	0.83	0.78	0.74		0.40	0.70	0.81	0.72		1.00	0.71
MDB 2002	0.84	0.87	0.81	0.69	0.73	0.80	0.67	0.71	0.68	0.64	0.77		0.75	0.62	0.70	0.08	0.71	1.00

1) Only correlations that relate to at least 6 countries are reported

Table 2:¹⁾
Kendall's
Rank
Correlation

	IMD 2003	IMD 2004	IMD 2002	PERC2002	PERC2003	PERC2004	GCR 2002	GCR 2003	GCR 2004	EIU 2004	FH 2004	TI/GI 2002	CU 2003	WMRC 2004	MIG	BEEPS02	Inf_Int2003	MDB 2002
IMD 2003	1.00	0.91	0.90	0.82	0.82	0.85	0.83	0.83	0.84	0.76	0.55	0.63	0.71	0.78	0.70	0.56	0.63	0.60
IMD 2004	0.91	1.00	0.89	0.72	0.77	0.80	0.81	0.83	0.82	0.78	0.47	0.62	0.71	0.79	0.73	0.61	0.55	0.60
IMD 2002	0.90	0.89	1.00	0.82	0.82	0.90	0.86	0.82	0.82	0.77	0.29	0.69	0.73	0.78	0.71	0.57	0.60	0.45
PERC2002	0.82	0.72	0.82	1.00	0.89	0.85	0.74	0.69	0.63	0.69		0.72	0.63	0.77	0.72		0.52	0.39
PERC2003	0.82	0.77	0.82	0.89	1.00	0.87	0.76	0.76	0.65	0.76		0.67	0.72	0.81	0.77		0.62	0.59
PERC2004	0.85	0.80	0.90	0.85	0.87	1.00	0.83	0.83	0.72	0.75		0.76	0.67	0.84	0.83		0.49	0.65
GCR 2002	0.83	0.81	0.86	0.74	0.76	0.83	1.00	0.83	0.81	0.72	0.28	0.68	0.66	0.73	0.68	0.46	0.59	0.44
GCR 2003	0.83	0.83	0.82	0.69	0.76	0.83	0.83	1.00	0.80	0.70	0.53	0.58	0.57	0.70	0.63	0.71	0.60	0.47
GCR 2004	0.84	0.82	0.82	0.63	0.65	0.72	0.81	0.80	1.00	0.69	0.72	0.59	0.62	0.71	0.63	0.48	0.62	0.44
EIU 2004	0.76	0.78	0.77	0.69	0.76	0.75	0.72	0.70	0.69	1.00	0.78	0.70	0.71	0.78	0.66	0.57	0.58	0.57
FH 2004	0.55	0.47	0.29				0.28	0.53	0.72	0.78	1.00		0.66	0.69	0.65	0.46		0.68
TI/GI 2002	0.63	0.62	0.69	0.72	0.67	0.76	0.68	0.58	0.59	0.70		1.00	0.73	0.69	0.65		0.33	
CU 2003	0.71	0.71	0.73	0.63	0.72	0.67	0.66	0.57	0.62	0.71	0.66	0.73	1.00	0.56	0.59	0.53	0.51	0.56
WMRC 2004	0.78	0.79	0.78	0.77	0.81	0.84	0.73	0.70	0.71	0.78	0.69	0.69	0.56	1.00	0.64	0.34	0.67	0.48
MIG	0.70	0.73	0.71	0.72	0.77	0.83	0.68	0.63	0.63	0.66	0.65	0.65	0.59	0.64	1.00	0.31	0.51	0.52
BEEPS02	0.56	0.61	0.57				0.46	0.71	0.48	0.57	0.46		0.53	0.34	0.31	1.00		0.15
Inf_Int2003	0.63	0.55	0.60	0.52	0.62	0.49	0.59	0.60	0.62	0.58		0.33	0.51	0.67	0.51		1.00	0.29
MDB 2002	0.60	0.60	0.45	0.39	0.59	0.65	0.44	0.47	0.44	0.57	0.68		0.56	0.48	0.52	0.15	0.29	1.00

1) Only correlations that relate to at least 6 countries are reported

These data refer to all countries, even those not included in the CPI.⁴ The correlations

on average are 0.81 for the Pearson correlation and 0.67 for Kendall's rank correlation. This suggests that the sources do not differ

⁴ Abbreviations are: BEEPS: Business Environment and Enterprise Performance Survey; CU: Columbia University; EIU: Economist Intelligence Unit; FH: Freedom House; II: Information International; IMD: Institute for Management Development; MDB: A Multinational

Development Bank; MIG: Merchant International Group; PERC: Political and Economic Risk Consultancy; TI/GI: Gallup International on behalf of Transparency International; WEF: World Economic Forum; WMRC: World Markets Research Centre.

considerably in their assessment of levels of corruption.

Confidence range

We have been providing the public with information on the confidence range for some years now. Up to 2001 these were based on the determination of the standard error for a country's average score and a resulting parametric assessment of a 95 confidence range. This approach required the assumption that there is no imprecision associated with the source's values and that these values are independent of each other. Another strong assumption required is that errors are normally distributed. While it is statistically difficult to relax the first two assumptions, one can relax the assumption of a normal distribution and apply tests that are valid for any type of distribution. Another drawback of the older confidence ranges was, again, that they sometimes violated the given range from 0 to 10. For example, while in 2001 Bangladesh had a score of 0.4, its 95% confidence range was between -3.6 and 4.4. For Finland, on the other hand, the range went as high as 10.4. This type of a range is confusing even to an expert. Since it is in contradiction to the official range reported, the public is equally disoriented.

In order to restrict the confidence range to our pre-specified limits we now apply a different methodology: a non-parametric approach applying the bootstrap methodology. The principal idea of such a bootstrap confidence range is to resample the sources of a country with replacement. If five source values (3, 5, 4, 4.5, 4.2) had been given, an example of such a sample would be (5, 5, 4.2, 3, 3). A sufficiently large number of such samples (in our case 10,000) are drawn from the available vector of sources and the sample mean is determined in each case. Based on the distribution of the resulting means, inferences on the underlying precision can be drawn. The lower (upper) bound of a 90% confidence range is then determined as the value where 5% of the sample's means are below

(above) this critical value.⁵ In addition to the "percentile" method just described, more complicated approaches exist. First, the confidence levels can be adjusted if (on average) the mean of a bootstrap sample is smaller than the observed mean. The relevant parameter is called z_0 . Another adjustment is to assume the standard deviation also to be dependent on the mean of the bootstrap sample. The relevant parameter is a . If both these adjustments are considered, the resulting approach is called a bootstrap-BC_a-method (bias-corrected-accelerated). A precise description of this approach can be obtained from Efron and Tibshirani (1993, chap. 14.3, 22.4 and 22.5).⁶ One concern with the BC_a approach is that it is throwing a lot of machinery at very few observations. Due to statistical considerations, a simple method might prove superior. Brad Efron had therefore suggested the use of a BC-approach for our purpose. In this case, z_0 is determined endogenously from the bootstrap sample but a is set equal to zero. There are two interesting characteristics of the resulting confidence range.

- 1) When requiring a 90% confidence range (which allows with 5% probability that the true value is below and with 5% probability that the value is above the determined confidence range) the upper (lower) bound will not be higher (lower) than the highest (lowest) value provided

⁵ There can arise boundary effects when only 3 or 4 sources exist. Only 10 different combinations are possible in the case of 3 sources, suggesting that a 5% confidence point can "hit" the boundary. If this is the case, the BC-approach could produce at random two different values for the upper (or the lower) confidence point. These boundary effects have been identified and, if existent, the more conservative range is reported in the table.

⁶ See Efron, B. and R. Tibshirani (1993), *An Introduction to the Bootstrap*, Chapman & Hall: New York and London: 202-219.

by a source. This implies that our range from 0 to 10 will never be violated.

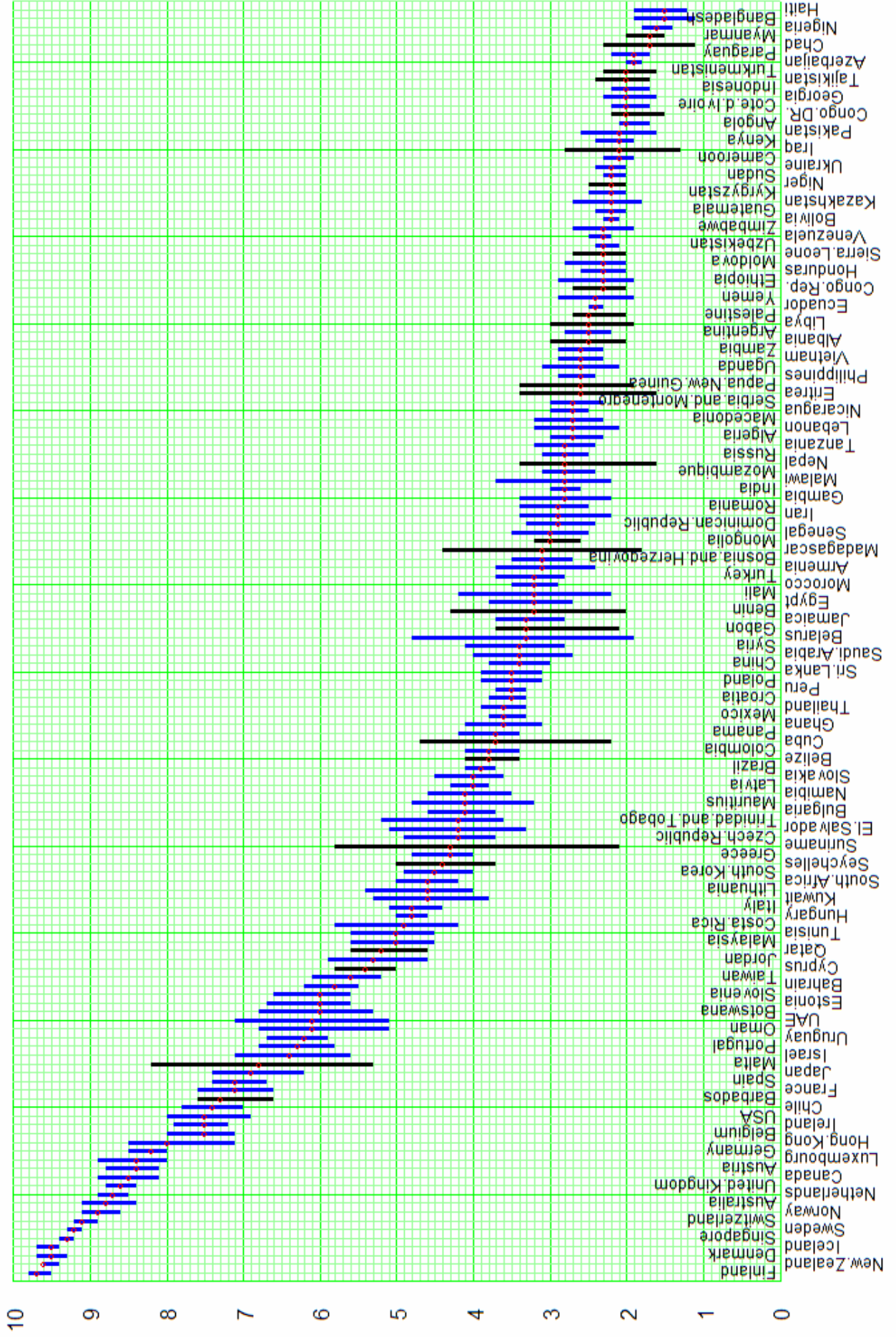
- 2) The confidence range remains valid even if the data (i.e. the standardized values for a given country) are not normally distributed. The range is even free of assumptions with regard to the distribution of these data.

However, with only few sources being used, there is a downward bias in the confidence range thus reported. When only few sources are available these do not fully capture the whole range of possible values. This misrepresentation becomes the larger the fewer sources are available. This issue is part of a general statistical problem that is not specific to our application: One simply cannot expect accurate estimates of a confidence interval from few observations.

In order to determine the size of this bias Walter Zucchini and Florian Hoffmann from the Institute for Statistics and Econometrics, University of Göttingen, wrote a short unpublished research paper. Given that the data are approximately beta distributed, various simulation tests were required. They found that the unbiased coverage probability is lower than its nominal value of 90%. The accuracy of the confidence interval estimates increases with a growing number of sources (n). The mean coverage probability is 65.3% for $n=3$; 73.6% for $n=4$; 78.4% for $n=5$; 80.2% for $n=6$ and 81.8% for $n=7$. While the confidence range nominally relates to a 90% level, an unbiased estimate of the confidence level is lower.

When interpreting the confidence range these results have to be born in mind. Figure 2 portrays the confidence ranges alongside with the scores.

Figure 2: 2004 CPI and approximate confidence intervals
The coverage probability is 65%-75% (black lines) or 80%-90% (blue lines)



Annex: Sources for the TI Corruption Perceptions Index (CPI) 2004

Number	1	2	3
Abbreviation	BEEPS	CU	EIU
Source	World Bank and the EBRD	Columbia University, The Center for International Earth Science Information Network	Economist Intelligence Unit
Name	Business Environment and Enterprise Performance Survey	State Capacity Survey	Country Risk Service and Country Forecast
Year	2002	2003	2004
Internet address	info.worldbank.org/governance/beeps2002/	http://www.ciesin.org/	www.eiu.com
Who was surveyed?	Senior businesspeople	US-resident country experts (policy analysts, academics and journalists)	Expert staff assessment
Subject asked	Frequency of irregular additional payments; how problematic is corruption for business?	Severity of corruption within the state	The misuse of public office for private (or political party) gain.
Number of replies	6,500	224	Not applicable
Coverage	25 transition countries	95 countries	142 countries

Number	4	5
Abbreviation	FH	II
Source	Freedom House	Information International
Name	Nations in Transit	Survey of Middle Eastern Businesspeople
Year	2004	2003
Internet address	www.freedomhouse.org/research/nattransit.htm www.freedomhouse.hu	www.information-international.com
Who was surveyed?	Assessment by US, regional, and in-country experts	Senior businesspeople from Bahrain, Lebanon and UAE
Subject asked	Extent of corruption as practiced in governments, as perceived by the public and as reported in the media as well as the implementation of anticorruption initiatives.	How common are bribes, how costly are they for doing business and how frequently are public contracts awarded to friends and relatives in neighboring countries
Number of replies	Not applicable	382 assessments from 165 respondents
Coverage	28 countries/territories	31 countries

Number	6	7	8
Abbreviation	IMD		
Source	International Institute for Management Development, Lausanne, Switzerland		
Name	World Competitiveness Yearbook		
Year	2002	2003	2004
Internet address	www.imd.ch		
Who was surveyed?	Executives in top and middle management; domestic and international companies		
Subject asked	Bribing and corruption exist in the economy		
Number of replies	3,532	> 4,000	4166
Coverage	49 countries	51 countries	

Number	9	10
Abbreviation	MDB	MIG
Source	A Multilateral Development Bank	Merchant International Group
Name	Survey	Grey Area Dynamics
Year	2002	2004
Internet address		www.merchantinternational.com
Who was surveyed?	Experts within the bank were identified and multiple questionnaires (each relating to a different country) were sent out to them. Roughly 40% of the questionnaires were returned.	Expert staff and network of local correspondents
Subject asked	How widespread is the incidence of corruption? (Widespread; Somewhat widespread; Somewhat limited; Limited; No judgment)	Corruption, ranging from bribery of government ministers to inducements payable to the "humblest clerk".
Number of replies	398	Not applicable
Coverage	47 countries	155 countries

Number	11	12	13
Abbreviation	PERC		
Source	Political & Economic Risk Consultancy		
Name	Asian Intelligence Newsletter		
Year	2002	2003	2004
Internet address	www.asiarisk.com/		
Who was surveyed?	Expatriate business executives		
Subject asked	How bad do you consider the problem of corruption to be in the country in which you are working as well as in your home country?		
Number of replies	More than 1,000	More than 1,000	More than 1,000
Coverage	14 countries		

Number	14	15
Abbreviation	TI/GI	WMRC
Source	Gallup International on behalf of Transparency International	World Markets Research Centre
Name	Corruption Survey	Risk Ratings
Year	2002	2004
Internet address	www.transparency.org/surveys/index.html#bpi	www.wmrc.com
Who was surveyed?	Senior businesspeople from 15 emerging market economies	Expert staff assessment
Subject asked	How common are bribes to politicians, senior civil servants, and judges and how significant of an obstacle are the costs associated with such payments for doing business?	The likelihood of encountering corrupt officials, ranging from petty bureaucratic corruption to grand political corruption.
Number of replies	835	Not applicable
Coverage	21 countries	186 countries

Number	16	17	18
Abbreviation	WEF		
Source	World Economic Forum		
Name	Global Competitiveness Report		
Year	2002	2003	2004
Internet address	www.weforum.org		
Who was surveyed?	Senior business leaders; domestic and international companies		
Subject asked	Undocumented extra payments connected with various government functions.		
Number of replies	Ca. 4,600	7,741	8,700
Coverage	76 countries	102 countries	104 countries