

## Research Reports



## MEASURING CORRUPTION: A COMPARISON BETWEEN THE TRANSPARENCY INTERNATIONAL'S CORRUPTION PERCEPTIONS INDEX AND THE WORLD BANK'S WORLDWIDE GOVERNANCE INDICATORS

ANJA ROHWER\*

### Introduction

Corruption is a complex social, political and economic phenomenon that is prevalent in all countries in varying degrees. There is no international consensus on the meaning of corruption. In the literature, corruption is commonly defined as the misuse of public power for private benefit (Lambsdorff 2007, 16). Although this definition has been widely adopted, several critics have observed that such a definition is culturally biased and excessively narrow (UNDP 2008, 12).

The crucial question is: is it possible to measure corruption, and if so, how? Corruption is a variable that cannot be measured directly. However, the number of indices focused on corruption measurement has grown exponentially over the past decade. They range from some of the more established and widely used indicators like Transparency International's (TI) Corruption Perceptions Index (CPI) and the World Bank's Worldwide Governance Indicators (WGI), to a newer generation of measurement and assessment tools like TI's Global Corruption Barometer and Global Integrity's Global Integrity Index. This article will discuss different ways to measure corruption and compare the different in-

dicators, focusing especially on CPI and WGI, over time. The CPI and the WGI are both composite indicators, made up of distinct component data sources that assess a wide and differing range of corruption (UNDP 2008, 6; Knack 2006, 15).

### The many faces of corruption

Corruption occurs basically in four main forms: bribery, embezzlement, fraud and extortion (Andvig et al. 2000, 14ff.):

- Bribery is understood as the payment (in money or kind) that is given or taken in a corrupt relationship. Equivalent terms to bribery include, for example, kickbacks, commercial arrangements or pay-offs. These are all notions of corruption in terms of the money or favours paid to employees in private enterprises, public officials and politicians. They are payments or returns needed or demanded to make things pass more swiftly, smoothly or more favourably through state or government bureaucracies.
- Embezzlement is theft of resources by people who are responsible for administering them, e.g., when disloyal employees steal from their employers. It is not considered corruption from a strictly legal point of view, but is included in a broader definition.
- Fraud is an economic crime that involves some kind of trickery, swindle or deceit. It involves manipulation or distortion of information, facts and expertise by public officials for their own profit.
- Extortion is money and other resources extracted by the use of coercion, violence or threats to use force.

Another popular way to differentiate various forms of corruption is by dividing it into petty and grand corruption. On the one hand, petty corruption is defined as street level, everyday corruption. It occurs when citizens interact with low- to mid-level public officials in places like hospitals, schools, police departments and other bureaucratic agencies. The scale of monetary transaction involved is small and primarily impacts individuals (and disproportion-

\* Ifo Institute for Economic Research at the University of Munich.  
E-mail: rohwer@ifo.de.

ably the poor; UNDP 2008, 8). On the other hand, grand corruption takes place at the policy formulation end of politics. It refers not so much to the amount of money involved as to the level at which it occurs (where policies and rules may be unjustly influenced). The kinds of transactions that attract grand corruption are usually large in scale. Grand corruption is sometimes used synonymously with political corruption (U4 – Anti-Corruption Resource Center, Glossary).

### Different methods for corruption measures

It is important to remember that there is no international consensus on the meaning of corruption. This ambiguity has direct implications on international rankings. And because of the many different faces of corruption and its very nature, it is almost impossible to deliver a precise and objective measure for the phenomenon. Objective data on corruption are difficult to obtain, and there is still no measurement system constructed that accurately accounts for actual levels of corruption within a country and, by extension, at the global level. This is because specific measures of corruption are imperfectly related to overall levels of corruption. In other words, current indicators are imperfect proxies for actual levels of corruption.

In recent years, however, corruption has been measured at the regional, national and global level, mostly using perception surveys as the leading method to collect data. Indicators have proved very useful in raising awareness, making cross-country comparisons and conducting statistical analysis, helping establish correlations between corruption and a wide range of variables (U4 – Anti-Corruption Resource Center 2009, 2).

We can distinguish between the following types of corruption indicators (UNDP 2008, 8 ff.):

- Perception-based indicators and experience-based indicators
- Indicators based on a single data source and composite indicators
- Proxy indicators.

Perception-based indicators are based on the opinions and perceptions of corruption in a given country among citizens and experts. Experience-based indicators measure citizens' or firms' actual experience

with corruption, i.e. whether they have been offered or given a bribe.

Indicators based on a single data source are produced by the publishing organisation without recourse to third-party data whereas composite indicators aggregate and synthesize different measures generated by various third-party data sources.<sup>1</sup>

Proxy indicators<sup>2</sup> try to measure corruption indirectly by aggregating as many opinions (or voices) and signals of corruption, or by measuring the opposite: anti-corruption, good governance and public accountability mechanisms.

The question now is what these indicators really measure. Because there are so many different forms of corruption, it is not possible for one indicator to capture the multidimensional aspect of corruption in a reliable and objective manner. Roughly speaking, all indicators of corruption are necessarily biased towards a specific dimension of corruption. For example, the World Bank's Control of Corruption indicator measures corruption in the public and private sector; hence, this indicator goes beyond the commonly accepted definition of corruption. The Transparency International's Corruption Perception Index measures corruption only in the public sector. To create an international corruption ranking system, it is essential to find an international consensus on what corruption is. Generally, no single indicator can capture the full complexity of the phenomenon. As a result it is more valuable to use a combination of tools rather than single indicators.

### Corruption indices

This section provides an overview of selected international corruption indices. We can distinguish between cross-country indicators of levels of corruption and cross-country assessments of anti-corruption frameworks. Both the Transparency International's Corruption Perceptions Index and the World Bank's Worldwide Governance Indicators (including Control of Corruption) belong to the former category. For more information about international corruption indices, see UNDP (2008), which gives an overview of more indices and tries to identify the different advantages of these indices.

<sup>1</sup> Composition indicators remain the most widely used measurement tools because of their near-global coverage.

<sup>2</sup> Relating to the belief that corruption is impossible to measure empirically.

*Transparency International's Corruption Perceptions Index (CPI)*

*Methodology*

The Transparency International's Corruption Perceptions Index ranks countries in terms of the degree to which corruption is perceived to exist among public officials and politicians. It is a composite index, a poll of polls, drawing on corruption-related data from expert and business surveys carried out by a variety of independent and reputable institutions. The CPI reflects views from around the world, including those of experts who are living in the countries evaluated.<sup>3</sup>

For the purpose of the CPI, which focuses on the corruption in the public sector, it is necessary to define corruption. TI defines "corruption as the abuse of public office for private gain" (Lambsdorff 2007, 16). The surveys used in compiling the CPI ask questions relating to the misuse of public power for private benefit. Examples include: bribery of public officials, kickbacks in public procurement, embezzlement of public funds and questions that probe the strength and effectiveness of anti-corruption efforts, thereby encompassing both the administrative and political aspects of corruption.

The CPI ranks countries on a zero to ten scale, with a score of zero representing very high corruption. Since each source of data about the level of corruption uses a different scale, the scores need to be standardized before being averaged into the CPI.<sup>4</sup> According to Lambsdorff (2007, 245 ff.) the standardization is carried out in two steps:

Until 2001 a simple mean and standard deviation approach was adopted for step 1.<sup>5</sup> The starting point for the standardization procedure was the previous year's CPI.<sup>6</sup> After standardization of each source, the simple average was taken for each country. Step 2 was a final standardization of the average values determined previously (Lambsdorff 2002, 6). A final standardization must be undertaken after aggregation of

the data because, due to the aggregation, the resulting mean and standard deviation can again differ from the previous year's values (Lambsdorff 1999, 9).

The new and currently used standardization procedure is carried out via a matching percentiles technique which determines the mean value for a country. This method uses the ranks of countries reported by each individual source. This method is useful for combining sources that have a different distribution.<sup>7</sup> And it allows all reported scores to remain in the bounds of CPI, i.e., to remain between zero and ten. The second step is a beta-transformation performed on scores to increase the spread of the distribution because averaging over several sources will tend to reduce the standard deviation in the sampling.<sup>8</sup>

Each country's score is then calculated as the average of all of the standardized values for the country. In order to create confidence intervals, TI uses a non-parametric bootstrap approach which allows inferences of the underlying precision of the results.<sup>9</sup> A 90 percent confidence interval is established where there is a 5 percent probability that the value is below and a 5 percent probability that the value is above this confidence interval. However, particularly when only few sources are available, an unbiased estimate of the mean coverage probability is lower than the nominal

<sup>6</sup> For example, in 2001 the 2000 CPI was the starting point for the standardization, except for older sources that were already standardized (where the standardized values determined in previous years were utilized). Standardization meant that the mean and standard deviation of a new source must take the same values as the respective subset in the 2000 CPI. The formula for

the standardized value  $S_i^j$

for the  $i^{\text{th}}$  component of CPI for the  $j^{\text{th}}$  country was then:

$$S_i^j = \left[ V_i^j - \mu_i^{sub} \right] * \frac{\sigma_{t-1}^{sub}}{\sigma_t^{sub}} + \mu_{t-1}^{sub}, \text{ where}$$

$V_i^j$  = the value of a particular corruption rating for a particular country in year t

$\mu_t^{sub}$  = mean of subgroup for a particular corruption rating in year t

$\sigma_{t-1}^{sub}$  = standard deviation of subgroup from CPI in year t-1

$\sigma_t^{sub}$  = standard deviation of subgroup from CPI in year t

$\mu_{t-1}^{sub}$  = mean of subgroup from year CPI in year t-1.

For more details see Lambsdorff 2002.

<sup>7</sup> For example, suppose that a particular data source on corruption for 2003 ranks three countries on their level of corruption. Then, in the first step of standardization, the best score (lowest corruption) from the 2002 CPI for those same three countries becomes the standardized score for the country with the highest ranking (lowest corruption) for that particular 2003 source. The second highest score among the 2002 CPI scores for the three countries is applied to the country with the second highest ranking according to the 2003 source, and so on (Thompson and Shah 2005, 4).

<sup>8</sup> For a more detailed description of the standardization procedure, especially the matching percentiles technique and beta transformation see Lambsdorff 2007.

<sup>9</sup> Confidence intervals are reported in order to underline the uncertainty with which corruption is measured.

<sup>3</sup> Hence, the index does not represent TI's own assessment.

<sup>4</sup> For example, one source rates countries on a 0-6 scale, while another rates on a 0-10 scale.

<sup>5</sup> Each of the sources had different means and standard deviations. Yet standardization does not mean that each source is given the same means and standard deviation, since each source covers a different subset of countries. Instead the aim of the standardization procedure is to ensure that inclusion of a source consisting of a certain subset of countries does not change the mean and standard deviation of this subset in the CPI. The reason is that the aim of each source is to assess countries relative to each other and not relative to countries not included in the source.

**Table 1**  
**CPI 1995–2008: Country coverage and sources**

Year	Number of countries included	Number of indicators	Number of independent sources
1995	41	7	3
1996	54	10	6
1997	52	7	6
1998	85	12	7
1999	99	17	10
2000	90	16	8
2001	91	14	7
2002	102	15	10
2003	133	17	13
2004	146	18	12
2005	159	16	10
2006	163	12	9
2007	180	14	12
2008	180	13	11

Source: Various documents available on TI's website. For a detailed list of all sources used in each year's CPI see [http://www.transparency.org/policy\\_research/surveys\\_indices/cpi](http://www.transparency.org/policy_research/surveys_indices/cpi).

value of 90 percent. Hence, the accuracy of the confidence interval estimates increases with the number of sources.

For generating the CPI it is necessary to have data from sources that span the last two years.<sup>10</sup> To provide a smoothing effect, TI includes 2 years data (if available) from surveys. In the case of experts' assessments only the most recent iteration of the assessment is included because these scores are considered to be the product of careful evaluation and change very little from year to year (Thompson and Shah 2005, 2).

Table 1 gives a review of changes in the number of sources used in the index as well as the number of countries scored. Only countries for which at least three sources are available are included in the CPI. Changes in the number of sources are driven by dropping outdated sources because only data which span the last two years are included in the CPI. The number of independent sources differs between countries. The sources used in the CPI 2008 are shown in Box 1.<sup>11</sup>

<sup>10</sup> For the CPI 2008, this includes surveys from 2008 and 2007.

<sup>11</sup> For an overview of all sources used in each year's CPI see TI's website.

<sup>12</sup> In 2008, for example, the data provided by the United Nations Economic Commission for Africa dropped out of the index. The data used in 2007 were no longer current and their new report is not yet available. As a result of the elimination of this source and other sources changing and expanding their coverage, some countries were affected by slight changes in the composition of sources (Lambsdorff 2008, 3). With respect to the methodology there was a change in the standardization procedure in 2002.

### A Year-to-Year Comparison

Year-to-year comparisons are difficult. A country's rank can easily change because new countries enter in the index and others drop out. Table 2 shows the results of the CPI from 1996 until 2008. Each year provides a snapshot of the views of business people and country analysts for the implied year. Comparisons with previous years should be made with caution because changes in a country's score can result from a changed perception of a country's performance or from a change in the CPI's sample and methodology.<sup>12</sup>

To compare a country's score over time it is necessary to go back to the individual sources.

### World Bank's Worldwide Governance Indicators: Control of Corruption

#### Methodology

The researchers at the World Bank adopted the basic approach of the TI CPI, but attempted to improve on it in several respects (Kaufmann et al. 1999) in their Worldwide Governance Indicators (WGI) project. This project reports aggregate and individual governance indicators for 212 countries and territories over

#### Box 1

##### Sources of the CPI 2008

###### Institutional Profiles Database

- **ADB:** Country Performance Assessment Ratings by the Asian Development Bank.
- **AFDB:** Country Policy and Institutional Assessment of the African Development Bank.
- **BTI:** Bertelsmann Transformation Index, Bertelsmann Foundation.
- **CPIA:** Country Policy and Institutional Assessment by the World Bank.
- **EIU:** Economist Intelligence Unit.
- **FH:** Freedom House Nations in Transit.
- **GI:** Global Insight (formerly World Markets Research Centre), Country Risk Ratings.
- **IMD:** International Institute for Management Development, Lausanne.
- **MIG:** Grey Area Dynamics Ratings by the Merchant International Group.
- **PERC:** Political and Economic Risk Consultancy, Hong Kong.
- **WEF:** World Economic Forum.

Source: Lambsdorff (2008).

the period 1996–2008, for six dimensions of governance: Voice and Accountability, Political Stability/ Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption in the 2009 edition.

The aggregate indicators combine the views of a large number of enterprise, citizen and expert survey respondents in industrial and developing countries. The individual data sources underlying the aggregate indicators are drawn from a diverse variety of survey institutes, think tanks, non-governmental organizations, and international organizations.

To produce one of the six components, for example Control of Corruption, the World Bank started by collecting existing perception indicators. They used only those perception indicators that contain useful information for assessing the quality of Control of Corruption in different countries from their point of view. In the 2009 edition of the WGI, the Control of Corruption indicator refers to 25 sources with 40 indicators.

The World Bank uses an Unobserved Component Model (UCM) to aggregate the various responses to six broad clusters.<sup>13</sup> This type of model is used to account for the fact that corruption itself is not observable and one can only approximate it by aggregating the scores from given indicators. More precisely, the aggregation of the different sources used takes place in five steps:

1. All indicators from the same source in the Control of Corruption cluster are combined into a single indicator. The World Bank uses a simple, un-weighted average of all the existing indicators they use

<sup>13</sup> For a detailed description on the construction of the WGI, see Kaufmann et al. (2009, 98 ff.) and Arndt and Oman (2006).

**Table 2**

**Corruption Perceptions Index, 1996–2008**

	1996	1998	2000	2002	2004	2006	2008
Austria	7.59	7.50	7.70	7.80	8.40	8.60	8.14
Belgium	6.84	5.40	6.10	7.10	7.50	7.30	7.28
Bulgaria		2.90	3.50	4.00	4.10	4.00	3.63
Cyprus					5.40	5.60	6.35
Czech Republic	5.37	4.80	4.30	3.70	4.20	4.80	5.24
Denmark	9.33	10.00	9.80	9.50	9.50	9.50	9.31
Estonia		5.70	5.70	5.60	6.00	6.70	6.60
Finland	9.05	9.60	10.00	9.70	9.70	9.60	9.04
France	6.96	6.70	6.70	6.30	7.10	7.40	6.93
Germany	8.27	7.90	7.60	7.30	8.20	8.00	7.85
Greece	5.01	4.90	4.90	4.20	4.30	4.40	4.66
Hungary	4.86	5.00	5.20	4.90	4.80	5.20	5.12
Ireland	8.45	8.20	7.20	6.90	7.50	7.40	7.71
Italy	3.42	4.60	4.60	5.20	4.80	4.90	4.81
Latvia		2.70	3.40	3.70	4.00	4.70	5.03
Lithuania			4.10	4.80	4.60	4.80	4.61
Luxembourg		8.70	8.60	9.00	8.40	8.60	8.33
Malta					6.80	6.40	5.81
Netherlands	8.71	9.00	8.90	9.00	8.70	8.70	8.90
Poland	5.57	4.60	4.10	4.00	3.50	3.70	4.62
Portugal	6.53	6.50	6.40	6.30	6.30	6.60	6.09
Romania		3.00	2.90	2.60	2.90	3.10	3.81
Slovak Republic		3.90	3.50	3.70	4.00	4.70	4.98
Slovenia			5.50	6.00	6.00	6.40	6.72
Spain	4.31	6.10	7.00	7.10	7.10	6.80	6.45
Sweden	9.08	9.50	9.40	9.30	9.20	9.20	9.31
United Kingdom	8.44	8.70	8.70	8.70	8.60	8.60	7.68
Croatia			3.70	3.80	3.50	3.40	4.36
Macedonia					2.70	2.70	3.62
Norway	8.87	9.00	9.10	8.50	8.90	8.80	7.90
Switzerland	8.76	8.90	8.60	8.50	9.10	9.10	8.97
Turkey	3.54	3.40	3.80	3.20	3.20	3.80	4.61
Australia	8.60	8.70	8.30	8.60	8.80	8.70	8.72
Canada	8.96	9.20	9.20	9.00	8.50	8.50	8.74
Japan	7.05	5.80	6.40	7.10	6.90	7.60	7.30
New Zealand	9.43	9.40	9.40	9.50	9.60	9.60	9.34
United States	7.66	7.50	7.80	7.70	7.50	7.30	7.28

Note: Only European and non-European OECD countries are included. Corruption Perceptions Index relates the perceptions of the degree of corruption as seen by business people and country analysts, and ranges between 10 (highly clean) and 0 (highly corrupt).

Source: Transparency International ([www.transparency.org](http://www.transparency.org)).

2. Now each of the constructed indicators is rescaled in order to determine whether the source covers a large enough number of countries in different income categories and regions to classify representative sources. For the composite indicator Control of Corruption, out of 25 sources used in 2009, nine are classified as representative.
3. The World Bank aggregates these representative sources into a preliminary composite indicator



**Table 3**  
**WGI Control of Corruption 1996–2008: country coverage and sources**

Year	Number of countries included	Median number of sources per country <sup>a)</sup>
1996	154	4
1998	194	5
2000	196	6
2002	197	7
2003	198	8
2004	205	9
2005	205	9
2006	207	11
2007	208	11
2008	208	11

<sup>a)</sup> Over time, there has been a steady increase in the number of sources included in the indicator. This increase in the number of data sources is reflected in an increase in the median number of sources available per country.

Source: Kaufmann et al. (2009, 32).

Control of Corruption. The different representative sources are weighted according to the strength of their correlation with one another. More precisely, the weight assigned to the numbers from each source is inversely proportional to its error variance relative to the other sources used to construct the indicator, so that the more closely the numbers from one source correlate with those of other sources, the lower the source’s perceived error variance is and the higher the weight is.<sup>14</sup>

4. The World Bank regresses the non-representative sources on the previously calculated composite indicator to obtain estimates of the error variances of these sources (i.e., the other 16 sources used to produce the Control of Corruption composite indicator). Here the World Bank uses the same weighting procedure as in step 3. The non-representative sources are used because they are corrected for the attenuation bias imparted by the measurement error in the estimates of the preliminary composite indicator in step 3 (Arndt and Oman 2006, 107).
5. Finally, new indicators draw on all indicators, representative and non-representative. Based on that, one can calculate new estimates for the true level of Control of Corruption for each country.

Kaufmann et al. (2009) construct each of their composite indicators in such a way that the resulting esti-

mates of governance have an expected value (across countries) of zero, and a standard deviation (across countries) of one.<sup>15</sup> This implies that virtually all scores lie between – 2.5 and 2.5, with higher scores corresponding to better outcomes.<sup>16</sup> This technical feature of the aggregate estimates also implies that their scales are largely arbitrary and that they cannot be reliably used for monitoring changes in levels of governance over time (Arndt and Oman 2006, 61), but according to Kaufmann et al. (2009, 15), they are informative about changes in individual countries’ relative positions over time.

The main advantage of this approach is that the aggregated indicators are more informative about unobserved governance than any individual source. Crucially the described method also generates margins of errors for the estimates of governance for each country, which need to be taken into account when making comparisons of governance across countries and over time (Kaufmann et al. 2009, 6).

Kaufmann et al. (2009) provide statistical confidence intervals<sup>17</sup> for each country’s score on each indicator in a given year. The difference between scores is meaningful only in the case of countries where scores differ by so much that their confidence intervals do not overlap (Arndt and Oman 2006, 29). The greater the number of sources the World Bank uses to generate the composite indicator for the country and the more closely these sources are correlated with each other, the smaller the country’s confidence interval (Arndt and Oman 2006, 64) will be.

In “Control of Corruption” the World Bank includes several indicators which measure the extent to which public power is exercised for private gain, including petty and grand forms of corruption, as well as “capture” of the state by elites and private interests. Despite this straightforward focus, the particular aspect of corruption measured by the various sources differs somewhat, ranging from the frequency of “additional payments to get things done,” to the effects of corruption on the business environment, to measuring “grand corruption” in the political arena, or to

<sup>14</sup> The logic of this weighting procedure is based on the assumption that errors in the numbers of one source are both independent across countries and independent of the errors in the numbers from the other sources used to construct the same opposite indicator, i.e. sources that tend to be highly correlated with other sources are assumed to be more informative.

<sup>15</sup> The estimates of Control of Corruption for each country are rescaled by subtracting the mean across countries and dividing by the standard deviation across countries.

<sup>16</sup> These boundaries correspond to the 0.005 and 0.995 percentiles of the standard normal distribution. For a handful of cases (for example Finland in 2006 and 2007 in Table 4), individual country ratings can exceed these boundaries when scores from individual data sources are particularly high or low.

<sup>17</sup> The confidence intervals are defined as the country’s score plus and minus 1.64 times its standard error.

**Box 2**

**Sources of the WGI Control of Corruption Indicator 1996–2008**

- **ADB:** Country Policy and Institutional Assessment of the African Development Bank.
- **AFR:** Afrobarometer.
- **ASD:** Country Performance Assessment Ratings by the Asian Development Bank.
- **BPS:** Business Environment and Enterprise Performance Survey.
- **BRI:** Business Environment Risk Intelligence.
- **BTI:** Bertelsmann Transformation Index, Bertelsmann Foundation.
- **DRI:** Global Insight Global Risk Service.
- **EIU:** Economist Intelligence Unit.
- **FRH, CCR:** Freedom House.
- **GCB:** Transparency International Global Corruption Barometer.
- **GCS:** World Economic Forum Global Competitiveness Survey.
- **GII:** Global Integrity Index.
- **GWP:** Gallup World Poll.
- **IFD:** IFAD Rural Sector Performance Assessments.
- **IPD:** Institutional Profiles Database.
- **LBO:** Latinobarometro.
- **MIG, GAD:** Cerberus Corporate Intelligence Gray Area Dynamics.
- **PIA:** World Bank Country Policy and Institutional Assessments.
- **PRC:** Political and Economic Risk Consultancy Corruption in Asia Survey.
- **PRS:** Political Risk Services International Country Risk Guide.
- **VAB:** Vanderbilt University Americas Barometer.
- **WCY:** Institute for Management Development World Competitiveness Yearbook.
- **WMO:** Global Insight Business Risk and Conditions.

Source: Kaufmann et al. (2009).

the tendency of elite forms to engage in “state capture”. The presence of corruption is often a manifestation of a lack of respect on the part of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules which govern their interactions, and hence represents a failure of governance according to the definition of the World Bank (Kaufmann et al. 2003, 4).

Table 3 gives a summary of some key features of the Control of Corruption indicator. Since 1996 there has been a steady increase in the number of countries and in the number of sources included in the WGI. Table 4

reports for a selection of countries the Control of Corruption indicator for the years 1996–2008. The sources used in the WGI 2008 are shown in Box 2.

*A Year-to-Year Comparison*

The aggregate WGI measures are not informative about trends in global averages. According to Kaufmann et al. (2009, 22) there is an assumption that world averages of governance are zero in each period, as a convenient choice of units. To assess trends in global averages of governance one has to return to the underlying individual data sources.

**Comparison of CPI and WGI**

The CPI and the WGI are both aggregate indicators which combine information from multiple sources. The WGI consists, as mentioned above, of six aggregate indicators while the CPI measures only corruption. In 2008 the CPI was calculated from a small set of data from 11 different organizations. The WGI Control of Corruption indicator used these 11 data sources from the CPI, as well as 14 others not used in the CPI.

A closer look at both indicators reveals that the WGI measures corruption in the public and private sector (with the help of some sources which provide data on corruption at the household level) as perceived by experts and opinion polls. The CPI measures corruption only in the public sector, as perceived by experts only.

Another distinction exists in the weighting scheme. The WGI weights available sources differently, in contrast to the equal weighting in the CPI of available sources for each country.<sup>18</sup>

The CPI contains statistical uncertainty and the WGI attempts to improve on it. While the CPI lists the number of sources and the range and standard deviation among sources, the WGI computes a standard error as an indicator of uncertainty accompanying each point estimate. These standard errors are lower for countries

- covered by more data sources and
- covered by data sources which are more highly correlated with other sources in the index (Knack 2006, 17).

<sup>18</sup> But there is an exemption for sources which have data available for more countries.

**Table 4**

**WGI Control of Corruption 1996–2008**

	1996	1998	2000	2002	2004	2006	2008
Austria	1.95	1.88	1.85	1.96	2.04	1.93	1.82
Belgium	1.36	1.38	1.52	1.60	1.47	1.33	1.35
Bulgaria	-0.80	-0.29	-0.14	-0.03	0.19	-0.04	-0.17
Cyprus	1.63	0.89	0.86	0.96	0.82	0.88	1.04
Czech Republic	0.66	0.49	0.30	0.36	0.40	0.36	0.37
Denmark	2.29	2.18	2.12	2.21	2.32	2.35	2.32
Estonia	-0.03	0.45	0.66	0.73	1.00	0.95	0.94
Finland	2.29	2.24	2.32	2.45	2.43	2.56	2.34
France	1.41	1.47	1.43	1.30	1.42	1.46	1.43
Germany	2.06	2.09	1.97	1.96	1.88	1.79	1.77
Greece	0.36	0.69	0.68	0.49	0.52	0.33	0.10
Hungary	0.66	0.72	0.77	0.64	0.75	0.62	0.55
Ireland	1.82	1.68	1.46	1.50	1.40	1.61	1.76
Italy	0.42	0.59	0.89	0.67	0.49	0.33	0.13
Latvia	-0.63	0.14	0.11	0.11	0.27	0.38	0.29
Lithuania	-0.15	0.27	0.42	0.31	0.39	0.20	0.18
Luxembourg	1.94	1.97	2.04	2.21	1.98	2.00	2.02
Malta	0.37	0.55	0.82	0.82	1.18	1.18	1.01
Netherlands	2.21	2.20	2.21	2.17	2.02	2.07	2.19
Poland	0.44	0.67	0.56	0.36	0.21	0.21	0.38
Portugal	1.56	1.30	1.19	1.31	1.19	1.06	1.08
Romania	-0.25	-0.36	-0.29	-0.36	-0.25	-0.14	-0.06
Slovak Republic	0.39	-0.02	0.31	0.13	0.49	0.41	0.43
Slovenia	1.10	1.02	0.85	0.87	1.07	1.01	0.95
Spain	1.04	1.38	1.40	1.39	1.38	1.14	1.18
Sweden	2.26	2.22	2.23	2.24	2.14	2.20	2.24
United Kingdom	2.19	2.17	2.10	2.06	1.95	1.87	1.77
Croatia	-0.57	-0.28	0.07	0.27	0.21	0.09	0.12
Macedonia	-1.06	-0.44	-0.54	-0.73	-0.44	-0.32	-0.11
Norway	2.28	2.18	2.05	2.09	1.95	2.07	1.88
Switzerland	2.18	2.17	2.10	2.14	2.06	2.18	2.15
Turkey	0.00	-0.22	-0.24	-0.46	-0.13	0.08	0.10
Australia	1.82	1.91	1.90	1.85	2.03	1.93	2.03
Canada	2.20	2.04	1.98	2.02	1.87	1.89	2.03
Japan	1.14	1.27	1.31	0.99	1.16	1.34	1.25
New Zealand	2.28	2.21	2.17	2.29	2.39	2.34	2.32
United States	1.72	1.70	1.73	1.84	1.73	1.29	1.55

Note: Only European and non-European OECD countries are included. The units in which the Control of Corruption is measured follow a normal distribution with a mean of zero and a standard deviation of one in each period. This implies that virtually all scores lie between -2.5 and 2.5, with higher scores corresponding to better outcomes.

Source: Kaufmann et al. (2009, 95ff.).

The motivation for developing aggregated indices is in both cases (CPI as well as WGI) the reduction of measurement errors by combining data from multiple sources.

**Criticism of composite indicators**

Measurement tools that aggregate a number of existing data sources, like the CPI or WGI, have their strengths and weaknesses. On the one hand, composite indicators can be useful in summarizing a lot of in-

formation from several sources, and in so doing they can limit the influence of measurement error in individual indicators and potentially increase the accuracy of measuring a concept as broad as corruption. On the other hand, one can run the risk of losing conceptual clarity.

*What is being measured?  
A sample bias and a transparency problem*

One problem in calculating aggregated corruption indices is the broad concept of corruption. It is unclear what the corruption indices actually tell us because the types of corruption and their meaning vary from one country to the next (Thompson and Shah 2008, 8). According to TI, the original purpose for their perceptions-based indicator was to raise awareness of corruption and to provide researchers with better data for analyzing the causes and consequences of corruption (Knack 2006, 16). The problem is that there are different forms of corruption. The degree of corruption in a country may depend on the frequency of corrupt acts, the amount of bribes paid or the gain that contractors achieve through corruption (Thompson and Shah 2005, 7). Besides the problem that it is not always clear what each of the different indicators used in the aggregated indices really measures, there is the problem that these different indicators measure different kinds of corruption. For example, the World Bank's Country Policy and Institutional Assessment asks about ineffective audits, conflicts of interest, policies being biased towards narrow interests, policies affected by corruption and public resources diverted to private gain whereas the World Economic Forum asks about the amount of bribes paid (Lambsdorff 2008). Therefore it is unclear what exactly the CPI and the WGI are measuring, when sources which measure such different aspects of corruption are averaged together.



Another problem is that both indices suffer from having definitions which can vary. The indices are dependent on the different sources which are responsive in the calculations. As shown in Table 1 and Table 3 the sources used in constructing these indices change over time. Hence the definitions of both indices change also. More precisely, the sources used in constructing the indices vary from country to country in a given year (Knack 2006, 18). For example, Latvia's 5.08 corruption rating and Lithuania's 4.61 corruption rating in the 2008 CPI (see Table 2) are based on two different sets of indicators, hence on varying implicit definitions of corruption. The same is true for the WGI Control of Corruption index 2008.

Another problem is the lack of transparency. This occurs on the one hand because some indicators rely on sources that are not publicly available (e.g., ICRG). But it happens that the aggregated indices are somewhat ambiguous because any component in the indices is constructed in an ambiguous manner (e.g., if the documentation in the ICGR provides little guidance as to how various aspects of corruption are weighted, or what information are used, it is impossible to fully explain, what the WGI Control of Corruption indicator is measuring or what the measurement is based on (Knack 2006, 17)).

#### *Whose perceptions are measured?*

As mentioned above, both indices, CPI as well as WGI, are based on perceptions of corruption. The data are collected from surveys and expert opinions that measures peoples' perception of corruption. Why? There are two reasons:

1. Objective criteria are hard to collect or they are too expensive for cross-country studies,
2. Available objective data is often misleading.

One way to avoid these problems is by asking people. By using carefully constructed surveys, perceptions measures about corruption may reflect realities of life better than objective measures. But perception-based indicators reflect perceptions. Therefore, their reflection of reality depends on whether perceptions reflect reality, and perceptions can change faster or more slowly than reality (Maurseth 2008, 27).

One of the strengths of CPI and WGI is that both indices contain sources based on the assessments of foreigners and sources based on samples of nationals. According to Galtung (2006, 112) the divergence of

the two groups of opinions are minimal and hence, "what counts as corruption in one part of the world, is understood similarly elsewhere".

To be a robust indicator for corruption it is necessary that the sample be homogeneous. The problem which occurs is that the different indicators used in the CPI and WGI are gathered from surveys based on questions to business people and very often the elite among business people. These business people are most likely representatives of multinational companies and represent views of only a small number of people. However, a good score for corruption on the company level does not result in a good score for corruption on the private level. This means that most influential factors like the experience of poor and disenfranchised people are ignored. Another problem is that most people are biased towards either a government or its opposition. Hence, objectivity is difficult to obtain (Thompson and Shah 2008, 12). Finally, every respondent's understanding of corruption is different depending on his or her experience and country of origin.

#### *Problems of using corruption as time series*

The standardization procedure used to place different indicators on a common scale excludes the ability to detect changes over time.<sup>19</sup> Not only the scores also the rankings are not comparable across years. The reason for this is that the composition of the samples changes.<sup>20</sup> Changes in the country's scores on CPI or WGI could be driven by adding a new source to the index or dropping an outdated one.<sup>21</sup> A possible solution would be to make a comparison for a single country at two points of time, if it were based on a common set of sources. However, it would be necessary to review the component data sources to see if there were changes in the methodology or in the definition over time.

#### *The likelihood of correlation of errors among sources: a specific WGI problem*

In the construction of the six composite indicators the World Bank uses a specific weighting scheme. This scheme is based on the assumption that the errors of the used sources are uncorrelated across sources and

<sup>19</sup> As mentioned in section of the methodology of the WGI the WGI is constructed to have a mean of 0 and a standard deviation of 1.

<sup>20</sup> i.e., the addition or purging of countries in the indices.

<sup>21</sup> For the CPI only data are included which span the last two years.

countries. This means, that sources which tend to be highly correlated with other sources are assumed to be more informative. Hence, the World Bank attributes higher weights to these sources than to sources that are less closely correlated with the majority. The sources which have low weights go into the composite indicator with a lower influence. This weighting scheme is very sensitive due to the composition of the composite indicator. For example, if five sources were to be aggregated in one composite indicator and four of them were very similar, the weights of the four similar sources will be overbalanced whereas the source which differs will have nearly no weight in the constructed composite indicator (Arndt and Oman 2006, 59).

Another problem in this context is that there are four reasons why the errors of the used sources are not uncorrelated across sources and countries (Arndt and Oman 2006, 66):

1. Experts who supply perception data which are used in one source by the World Bank are often informed of and influenced by the assessments of experts supplying such data for other sources in the WGI.
2. Experts who supply perception data that are used in the WGI for different sources are often informed of and influenced by perceptions and assessments from the same third parties (non-WGI source).
3. Crises and perceived changes or longer-term trends in a country's economic performance (like FDI inflows) often influences the perceptions used as inputs in the WGI.
4. Because the interpretation of survey questions is context- and cultural-specific, perception errors of different sources that rely on respondents from the same country or culture tend to be correlated.

Nevertheless, aggregated sources provide more information as compared to using only one source.

#### *The standardization and weighting procedure of the CPI*

The second phase of the standardization procedure based on the matching percentiles technique is advantageous in terms of data presentation. The scores will always fall between zero and ten. But this method has a defect: according to Galtung (2006, 120), it is natural to compare a country with its major regional neighbours in order to have a baseline of comparison

when countries are ranked by respondents or even expert panels. Even if a country tries to constrain corruption with reforms and it is assessed as being on a positive trend by some of the CPI's sources, its rank order in comparison to neighbouring countries may take years to change.<sup>22</sup> And it is getting harder if even other countries in a given region try the same. Hence, the matching percentiles technique punishes small improvements in a country's rank.

The CPI weights each of its sources equally with one exception: in 2008 the two most recent IMD and PERC surveys are each included as a separate source. Therefore each receive double weighting given to another source, such as FH and GI.

#### *A solution of weighting problems in the CPI and WGI*

Given that both corruption indices (CPI and WGI) have problems with their weighting schemes, the question arises what would be a plausible weighting scheme? According to Knack (2006, 27) a weighting scheme should ensure that:

- those sources that represent truly independent assessments are weighted more heavily,
- those sources with more extensive publicly available documentation and detailed justifications are weighted more heavily and
- those sources with larger and more nationally representative samples and those that include more questions on corruption are weighted more heavily.

The problem that still exists is that weights are practically always subjectively determined.

#### **Conclusion**

Generally speaking, corruption can be measured. The question remains as to how accurately. The analysis of two perception-based indices, the CPI and the WGI, resulted in some points of criticism. In conclusion such indices should be used with more caution because of lack of transparency and definition problems. There should be more control of the criteria and of the methods of obtaining aggregated indicators to better understand what they are measuring, and to determine (roughly) their degree of interdependence.

<sup>22</sup>This argument is based on comparisons of ranks and not of scores.

All users of the aggregated indices and their perceptions-based components should follow TI's example and acknowledge that these are measures of corruption perceptions, not of corruption. It should be remembered that perception-based indicators are not completely reliable because of the problems mentioned above.

The growth in the demand for, and use of, corruption indices have generated important new initiatives. The real value-added of these new initiatives is "that they produce indicators which are both transparent and concrete enough to be more directly useful for policy makers...to identify specific kinds of change" (Arndt and Oman 2006, 93). One example is the TI's Global Corruption Barometer, which assesses the general public's perception and experience of corruption. Two important features are that TI provides the disaggregated data free of charge and that the results are comparable across countries and over time (Transparency International, 2009, [http://www.transparency.org/policy\\_research/surveys\\_indices/gcb](http://www.transparency.org/policy_research/surveys_indices/gcb)).

## References

- Andvig, J. C., O.-H. Fjeldstad, I. Amundsen, T. Sissener and T. Søreide (2000), *Research on Corruption: A Policy Oriented Survey*, mimeo.
- Arndt, C. and C. Oman (2006), *Uses and Abuses of Governance Indicators*, OECD Development Centre Studies.
- Galtung, F. (2006), "Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices", in C. A. Sampford, C., A. Shacklock, C. Connors and F. Galtung, eds., *Measuring Corruption*, Ashgate, Hampshire, UK and Vermont, US, 101–130.
- Kaufmann, D. and A. Kraay, "Governance Indicators: Where Are We, Where Should We Be Going", *World Bank Policy Research Working Paper* 4370.
- Kaufmann, D., A. Kraay and M. Mastruzzi (2003), "Governance Matters III: Governance Indicators for 1996–2002", *World Bank Policy Research Working Paper* 3106.
- Kaufmann, D., A. Kraay and M. Mastruzzi (2009), "Governance Matters VIII: Aggregate and Individual Governance Indicators 1996–2008", *World Bank Policy Research Working Paper* 4978.
- Kaufmann, D., A. Kraay and P. Zoido-Lobaton (1999), "Aggregating Governance Indicators", *World Bank Policy Research Working Paper* 2195.
- Knack, S. (2006), "Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators", *World Bank Policy Research Working Paper* 3968.
- Lambsdorff, J. G. (1999), *The Transparency International Corruption Perceptions Index 1999 – Framework Document*, s\_indices/cpi/previous\_cpi/1999 (accessed 31 July 2009).
- Lambsdorff, J. G. (2002), *Background Paper to the 2002 Corruption Perceptions Index – Framework Document 2002*, s\_indices/cpi/2002 (accessed 31 July 2009).
- Lambsdorff, J. G. (2007), *The Institutional Economics of Corruption and Reform: Theory, Evidence, and Policy*, Cambridge University Press.
- Lambsdorff, J. G. (2008), *The Methodology of the Corruption Perceptions Index 2008*, [http://www.transparency.org/policy\\_research/surveys\\_indices/cpi/2008](http://www.transparency.org/policy_research/surveys_indices/cpi/2008) (accessed 31 July 2009).
- Maurseth, P. B. (2008), "Governance Indicators: A Guided Tour", *NUPI Working Paper* 754.
- Thompson, T. and A. Shah (2005), *Transparency International's Corruption Perceptions Index: Whose Perceptions Are They Anywhere?*, mimeo.
- Transparency International (2009), [http://www.transparency.org/policy\\_research/surveys\\_indices/gcb](http://www.transparency.org/policy_research/surveys_indices/gcb) (accessed 31 July 2009).
- U4 – Anti-Corruption Resource Centre (2009), *Measuring International Trends in Corruption*, <http://www.u4.no/helpdesk/helpdesk/query.cfm?id=199> (accessed 31 July 2009).
- U4 – Anti-Corruption Resource Center (2009), *Glossary*, <http://www.u4.no/document/glossary.cfm> (accessed 31 July 2009).
- UNDP (2008), *A Users' Guide to Measuring Corruption*, <http://www.globalintegrity.org/toolkits/books.cfm> (accessed 31 July 2009).