

The Transparency International

Corruption Perceptions Index 1999

— Framework Document

Dr. Johann Graf Lambsdorff

on behalf of Transparency International,

October 1999

1. The Methodology

1.1 Introduction

Transparency International's Corruption Perceptions Index (CPI) has assumed a central place in debates about corruption. It is used by economists, academics, business people and journalists. The growing importance of the CPI has stimulated interest in the methods used to compile it each year. This document, complementing the publication of the 1999 CPI and the press materials published with it, provides an in-depth explanation of the methodology.

The goal of the CPI is to provide data on extensive perceptions of corruption within countries. This is a means of enhancing understanding of levels of corruption from one country to another. It does not attempt to assess the degree of corruption practiced by nationals outside their own countries. This is a separate phenomenon and a separate instrument, the Bribe Payers Propensity Index (BPI), is published this year for the first time.

In an area as complex and controversial as corruption, no single source or polling method has yet been developed that combines a perfect sampling frame, large enough country coverage, and a fully convincing methodology to produce comparative assessments. This is why the CPI has adopted the approach of a composite index. It consists of credible surveys using different sampling frames and various methodologies and is the most statistically robust means of measuring perceptions of corruption.

1.2 Objective vs. Subjective Data

Unbiased, hard data is difficult to obtain and usually raises difficult questions with respect to validity. One such set of data has been assembled by the Crime Prevention and Criminal Justice Division of the United Nations Office at Vienna, [United Nations 1999]. This is a survey of national agencies in a large variety of countries called the United Nations Survey of Crime Trends and Operations of Criminal Justice Systems. The major goal of this investigation has been to collect cross-nationally comparative data on the incidence of reported crime and the operations of criminal justice systems. The questionnaire consists of a series of questions which ask for data, primarily statistical, on the main components of the criminal justice system. The latest version of this survey relates to the years 1990 to 1994. All national data are derived from the official national criminal statistics. Efforts have been undertaken to help national agencies in the collection of data and to review all parts of the survey for consistency.¹ However, the precise legal definition of bribery and corruption can be different in each national context, the differences drawn between bribery, embezzlement and fraud may be troublesome and the statistical methodology of counting and aggregating used in each national agency can differ considerably from that used elsewhere. Apart from this, countries such as Singapore and Hong Kong have extremely high per capita

¹ A full description of the methodology and the complete data can be obtained via internet at: <http://www.ifs.univie.ac.at/~uncjin/wcs.html>.

conviction rates for bribery. This lends itself to the conclusion that the data are to a large extent determined by the effectiveness and capacity of a country's judiciary in prosecuting corruption. High levels in this case indicate the success of anti-corruption initiatives rather than high levels of actual corruption. As such problems commonly arise with objective data, international surveys on perceptions serve as the most credible means of compiling a ranking of nations.

1.3 Sources in 1999

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 aimed at making the final decision as transparent and robust as possible. As a results of this it was decided that the 1999 CPI includes data from the following sources:

- ◆ Freedom House Nations in Transit (FH),
- ◆ Gallup International (GI),
- ◆ the Economist Intelligence Unit (EIU),
- ◆ the Institute for Management Development, Lausanne (IMD),
- ◆ the International Crime Victim Survey (ICVS),
- ◆ the Political and Economic Risk Consultancy, Hong Kong (PERC),
- ◆ The Wall Street Journal, Central European Economic Review (CEER),
- ◆ the World Bank and University of Basel (WB/UB),
- ◆ the World Economic Forum (WEF).

A number of other possible sources has been considered for inclusion and rejected because they mixed corruption with other issues such as xenophobia, nationalism or related to changing

corruption rather than levels of corruption. The 1999 CPI combines assessments from the past three years to reduce abrupt variations in scoring. Such changes might be due to high-level political scandals that affect perceptions, but do not reflect actual changing levels of corruption. For GI and WB/UB only one survey was conducted and this methodology cannot be applied. Also older data by CEER and ICVS cannot be used for this purpose because they are more than three years old. This suggests to adopt this methodology, as in the past, only for the surveys conducted by PERC, WEF and IMD. While this averaging is valuable for the inclusion of surveys, it is inappropriate for application to the data compiled by country experts. Such assessments as compiled by PRS, FH and EIU 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 judgement 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 elite or general public surveys an averaging over various years produces a useful smoothing effect. While some annual data may contain random errors, these do not necessarily carry over into the next year, and their impact is decreased by the averaging procedure. Alongside their annual report, the WEF in 1998 conducted a survey in various African countries which could also be included. Overall, 17 sources could be included into the 1999 CPI, originating from 10 independent institutions. The complete list of sources is presented in appendix 1.

The sources in the table indicate the countries to be included. The number of sources has increased further in 1999 as compared to 1998. This has allowed the number of countries in the CPI to be increased to 99. The increase in the number of sources, on the other hand, has allowed a high level of reliability to be maintained and partly increased.

1.4 Year-to-Year Comparisons

The CPI incorporates as many reliable and up-to-date sources as possible. One of the drawbacks to this approach is that 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 changes in sample and methodology. This is comparable to the problem of designing a price index for a basket of goods when the ingredients are constantly changing. The price index for one period cannot be fully compared to that of the next since the basket on which it is based has changed. A similar problem can arise with the CPI. Some sources are not updated and must be dropped as a result, while new, reliable sources are added. With differing respondents and slightly differing methodologies a change in a country's score cannot be attributed solely to actual changes in a country's performance.

Accordingly, TI repeatedly stresses that each year's index must be seen as the result of the sum of all reputable sources available at that time. Comparisons with the views collected in previous years can be misleading. In order to reduce the number of misleading interpretations of the CPI scores, the official CPI table will not include the scores from the previous year. Overall, however, the sources continue to show a high degree of correlation. So in practice, the impact of changes in samples and methodologies on the outcome appears to be rather small.

2. Validity

2.1 Defining Corruption

All sources generally apply a definition of corruption such as the misuse of public power for private benefits, e.g., 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:

- ◆ The IMD asks respondents to assess whether “Improper Practices (such as bribing and corruption) prevail or do not prevail in the public sphere.”
- ◆ The WEF asks “Irregular, additional payments connected with import and export permits, business licenses, exchange controls, tax assessments, police protection or loan application are common/ not common.”
- ◆ The PERC asks “To what extent does corruption exist in the country in which you are posted in a way that detracts from the business environment for foreign companies?”
- ◆ GI asks “From the following groups of people, can you tell me for each of them, if there are a lot of cases of corruption given, many cases of corruption, few cases or no cases of corruption at all.” The following groups were considered for the CPI: politicians, public officials, policemen and judges.
- ◆ The EIU defines corruption as the misuse of public office for personal (or party political) financial gain and aims at measuring the pervasiveness of corruption. Corruption is one of over 60 indicators used to measure “country risk” and “forecasting.”
- ◆ The ICVS asks: "In some areas there is a problem of corruption among government or public officials. During 1995, has any government official, for instance a customs officer, police officer or inspector in your own country, asked you or expected you to pay a bribe for his service?"
- ◆ The PRS determines a variable "Corruption in Government" and assesses the overall spread of corruption.
- ◆ FH and CEER determine the "level of Corruption" without providing further defining statements.
- ◆ The WB/UB asks two related questions with respect to corruption: first, “Please judge on a six point scale how problematic [corruption is] for doing business”; second, “It is common

for firms in my line of business to have to pay some irregular ‘additional payments’ to get things done. This is true always, mostly, frequently, sometimes, seldom or never.”

The terms "level", "problem", "prevalence", "pervasiveness", "commonness" and "number of cases" are largely identical. They all refer to some kind of “degree” of corruption, which is the 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, in none of these surveys have special forms of corruption been found to dominate or be more or less important than in others. The sources can be said to aim at measuring the same phenomenon.

It is important to note that largely none of the sources differentiates between administrative and political corruption, and that both types of corruption are addressed equally by the various questions posed. The IMD asks about corruption in the public sphere. This inevitably includes both corruption in administration and in politics, as they both constitute the public sphere. The WEF in 1997 and 1998 addressed only particular areas where corruption occur and in each of these, either politicians or administrators can be the relevant actors. Political corruption, like administrative corruption, requires additional payments and may represent an obstacle to doing business - the topic of the WB/UB. Similarly, corruption in government, as assessed by the PRS, also includes both types of corruption, since administration and politics are each parts of government structures. The same applies to the broad definitions used by FH and CEER. The EIU explicitly notes that its assessments include corruption among public servants and politicians alike. The GI data provides the only differentiation between political and administrative corruption. It is noteworthy in this respect to report a correlation of 0.88 between the assessment of politicians and that of administrators (an average of the assessment of judges, policemen and public servants) in their

data, pointing to a high correlation between the two aspects of the corruption phenomenon. This further justifies a blending of political and administrative corruption, since there is no strong evidence that countries differ in the prevalence of the one type of corruption over another. The only source which does not include large scale political corruption is the ICVS. But taking into account that this source well correlates with the other sources, there was no strong argument that also the extent of political corruption is not well represented by this data.

2.2 Degrees of Corruption

As we have emphasized, the CPI aims to assess the "degree of corruption". But this term can suggest different interpretations, Rose-Ackerman [1999: 4]. In order to confirm the validity of our approach, we must first clarify whether this term is unambiguous. Imagine the simple case that 10 percent of all public servants take a bribe of \$200 each, 5 times a year in exchange for awarding a contract that results in a gain of \$500 each for corrupt private contractors. What would be the "degree of corruption" in this case? Three different potential definitions of the "degree of corruption" might be suggested:

1. The frequency of corrupt acts (e.g. 10 percent of public servants are bribed 5 times a year, or 5 times 10%) could be the crucial variable to assess the overall degree of corruption.
2. The amount of bribes paid (e.g. \$200 times 5 times 10 %) could be assumed to be the adequate measure.
3. The overall gain that contractors achieve via corruption (e.g. \$5000 times 5 times 10 percent), thus calculated, could be said to best reflect the degree of corruption.

While all of these definitions appear to be valid, they need not fully correlate with each other. For example, consider that a few high-ranking public servants are taking large bribes on the one hand, as opposed to many public servants engaging in petty corruption on the other. The total sum of

bribes would be about the same in both cases, but the frequency of corrupt incidents would doubtlessly be higher in the latter case. Similarly, when corrupt private competitors are in a strong bargaining position and do not allow much of their illegitimate gain to be shared with public servants, the degree of corruption as defined by the total amount of bribes may be low while total gains from corruption, i.e. the third definition for the degree of corruption, may be large. On the other hand, "absence of corruption" would be similarly defined in all three cases — i.e. frequency, amount of bribes and value of rents — as being equal to or nearly zero.

Having taken this theoretical look at degrees of corruption, we can now turn to the particular definitions used by our sources. First, it is interesting to see that the GI data clearly refers to the number of cases of corruption, which is best contained in the first definition. Similarly, the questions asked by the ICVS, WEF and the WB/UB (second question) relates to the frequency of bribes paid, which is also closer to the first definition. In contrast to this, the questions by the PERC and the WB/UB (first question) hint at the damage to private business people caused by corruption. The implication here might be that large bribes are particularly serious, while large benefits for corrupt private business people in the sense of definition 3 may not be. This may relate to a definition of corruption according to definition 2. The questions posed by the IMD, PRS, CEER, FH and EIU provide no insight regarding an assessment of degree. The terms "level", "prevalence", "existence" and "pervasiveness" used there might refer to frequency as well as the overall value of bribes involved. In sum, either definition 1 or 2 might be applicable to the definition of the degree of corruption.

While there is no clear answer whether to choose definition 1 or 2, we must note here that the sources correlate well with each other. This allows us to argue that at the moment there is little evidence that differences with respect to these two definitions are crucial to the outcome of a survey. Either, respondents have a very homogeneous pre-specified idea of how to define the "degree of corruption" which influences their response, irrespective of the precise wording of the

questionnaire, or countries do not differ considerably with respect to the particular kinds of corruption that prevail there. The latter alternative is particularly interesting. Anecdotal evidence gives the impression that grand corruption is sometimes more prevalent in one country, while petty corruption may be dominant in another. More research is required to deepen our understanding of the levels and types of corruption and the extent to which corruption differs between countries. Yet whatever definition of corruption applied by the respondents, it does not appear to be crucial to the data we obtained and the index remains valid despite the fact that the term "degree of corruption" is not exhaustively defined.

Interesting is one question posed by the WEF in 1999 concerning government favoritism. Respondents were asked for the extent of their agreement to the following statement: "Sweetheart deals between well-connected private firms and the government are not common." One should assume that the resulting index is related to the third definition of corruption provided above, i.e. that large gains for private contractors can be obtained for small favors to public decision makers. It is interesting that the resulting index correlates 0.79 with the other index by the WEF included into the CPI. While this value is certainly large, it may appear that differences between indices relating to the value of rents as opposed that of bribes can be established at some time in the future.

2.3 Reliability

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. This approach has been widely endorsed, see e.g. Lancaster and Montinola [1997]. 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 from 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 1998 CPI can be drawn from the high correlation between the sources. As most correlations are around 0.8 or higher, the sources do not differ considerably in their assessment of levels of corruption. This data is in table 1. In addition to these correlations, the reliability for each individual country score can be determined. The table reporting on the results for each country includes data on the number of sources that included a country and the standard deviation between the sources. The larger the number of sources and the

Table 1: Correlations between Sources

	IMD-1999	IMD-1998	IMD-1997	PRS-1999	GI-1997	PERC-1999	PERC-1998	PERC-1997	WEF 1997	WEF 1998	WEF/Africa-98	WEF 1999	WB/UB-1997	EIU-1999	ICVS-1996/7	CEER, 1997
IMD-1998	0.97	1,00														
IMD-1997	0.97	0.98	1,00													
PRS-1998	0.70	0.70	0.69	1,00												
GI-1997	0.86	0.89	0.89	0.66	1,00											
PERC-1999	0.91	0.83	0.88	0.40	0.78	1,00										
PERC-1998	0.97	0.95	0.96	0.34	0.92	0.90	1,00									
PERC-1997	0.96	0.92	0.95	0.37	0.88	0.97	0.96	1,00								
WEF 1997	0.92	0.89	0.93	0.56	0.79	0.85	0.94	0.89	1,00							
WEFb 1998	0.92	0.87	0.91	0.65	0.77	0.86	0.90	0.87	0.94	1,00						
WEF/Africa-98	NA	NA	NA	0.49	NA	NA	NA	NA	NA	NA	1,00					
WEF 1999	0.90	0.83	0.88	0.58	0.77	0.84	0.93	0.87	0.93	0.96	NA	1,00				
WB/UB-1997	0.92	0.93	0.93	0.44	0.79	0.89	0.97	0.94	0.85	0.87	0.70	0.87	1,00			
EIU-1999	0.86	0.85	0.88	0.67	0.76	0.88	0.95	0.92	0.86	0.89	0.71	0.84	0.82	1,00		
ICVS-1996/7	0.91	0.92	0.92	0.76	0.93	NA	NA	NA	0.90	0.90	NA	0.89	0.91	0.86	1,00	
CEER, 1997	0.75	NA	NA	0.64	0.79	NA	NA	NA	0.43	0.95	NA	0.86	0.56	0.76	0.71	1,00
Freedom House	0.73	NA	NA	0.66	0.70	NA	NA	NA	0.33	0.86	NA	0.87	0.58	0.79	0.67	0.93

"NA" indicates that few (4 or less) common countries were available so that a correlation would not be meaningful.

A grey colour indicates that no more than 6 common countries were available for the reported correlation coefficient.

lower the standard deviation between the sources, the more reliable is the value for a country. The relatively large standard deviation for Bolivia of 1.1 signifies that 66% of the sources ranged between a value of 1.4 and 3.6. In contrast, the low standard deviation for Germany means that 66% of the scores range between 7.5 and 8.5. A measure of precision for the countries' scores can be derived, assuming that the values provided by the scores are not stochastic and are all independent from each other — assumptions which are not necessarily realistic and should rather be described as a first-best scenario. Dividing the standard deviation by the square root of the number of sources minus one yields a measure which represents the standard error of the mean score. This is an imperfect but still helpful measure of precision for the individual country scores. Germany, with 10 sources shows a value of 0.17, while a value of 0.49 is obtained for Bolivia. This formula illustrates that precision increases with the number of sources and decreases with the standard deviation.

Difficulties in assessing the level of corruption in a country may, on the one hand, result from subjective difficulties, for example lack of data or experience. On the other hand, objective difficulties may contribute to this. Assessing the overall level of corruption may be difficult in countries where some institutions openly engage in corruption while others strongly resist and fight corruption. A large standard deviation may in this respect also reflect a heterogeneous state of a society.

3. Perception and Reality

3.1 Sample Design

While the sources all aim at measuring the degree of corruption, the sample design differs considerably. With the exception of the PERC, FH, CEER, EIU and PRS the sources mostly sample residents, who must rely on their personal, local estimate (as opposed to an expatriate's external viewpoint) of the degree of corruption and the meaning of the term in their own cultural

context. Whether this difference may lead to different outcomes still requires scientific study. For the purposes of the CPI, it added to the robustness of the resulting figures, since the data correlate well with other data. This correlation suggests that there being different samples makes no great difference to the results.

Of greater importance is the difference between polls on the opinions of experts and the polls of the general public provided by GI and the ICVS. Whereas the general public may tend to form views on the corruption (or lack of it) experienced in daily life, business people and risk analysts are usually closer to high-level incidents of corruption and may be in a better position to assess grand corruption. Furthermore, elites may have a biased viewpoint towards corruption insofar as they might be less negative about forms of corruption which favor their own group. Similarly, the general public may be less negative about petty forms of corruption. To what extent the general public deviates from an elite sample in its assessment of corruption has not yet been the subject of investigation and constitutes an important area for future research. The relatively high correlation (mostly around 0.8 or higher) of the GI and ICVS data with the other sources suggests no significant difference between these viewpoints.²

Each country reported in the CPI has been included in at least one of the surveys by GI, ICVS, IMD, WEF or WB/UB. This means that no country is assessed by expatriates only. Local residents have contributed to the assessment of all of the countries included in 1999. Overall, the 1999 CPI contains the perceptions of nationals from at least 99 countries. These are people from all continents and regions of the world.

² Even when elite and general public viewpoints show some differences, an aggregation of these data still makes sense, just like price levels for various goods can be aggregated to form some combined price index. While the idea of creating a price index would be to value a complete basket of goods, the idea of aggregating subjective data would then be to obtain an assessment of the level of corruption as seen by a broad and possibly heterogeneous sample of respondents.

3.2 Interpreting Perceptions

As the data collected relates to perceptions rather than to real phenomena, it has to be considered whether such perceptions improve our understanding of what real levels of corruption may be. This is necessary for the CPI to be a fruitful contribution to political debate, investment decisions and academic research. Since actual levels of corruption cannot be determined directly, perceptions may be all we have to guide us. However, this approach is undermined to at least some extent, if the perceptions gathered are biased. Such a potential bias might originate from the particular cultural background of respondents. This has been remarked by Bayley, [1966: 721]:

"The western observer is faced with an uncomfortable choice. He can adhere to the Western definition, in which case he lays himself open to the charge of being censorious and he finds that he is condemning not aberrant behavior but normal, acceptable operating procedure.... On the other hand, he may face up to the fact that corruption, if it requires moral censure, is culturally conditioned ... [and] it may be necessary then to assert in the same breath that an official accepts gratuities but is not corrupt."

In determining the level of corruption, the Bayley's viewpoint assigns a much more active role to the attitudes towards political and administrative behavior. Perceptions would not allow conclusions to be drawn with respect to real levels of corruption. Depending on whether the sample consist of locals or expatriates, this suggests two potential biases to be relevant.

Imagine that being asked to assess the level of corruption, a local estimates a high level of corruption in the country of residence. Such an assessment would be a valid contribution to the CPI only if the respondent makes the assessment as a result of comparisons with the levels of corruption perceived in other countries. But this is not necessarily the viewpoint taken by the

respondent. A respondent may also assign high levels by comparing corruption to other (potentially less pressing) problems facing the country, or by evaluating it according to a high ethical standard (e.g. which assumes any kind of gift-giving to a public official to be corrupt and not culturally acceptable). In the case of such an outlook, a high degree of observed corruption may reflect a high standard of ethics rather than a high degree of real misbehavior. Perceptions would be a misleading indicator for real levels of corruption. This bias can occur particularly if only locals are surveyed, each assessing only the levels of perceived corruption in their own countries. If respondents are asked to assess foreign countries or to make comparisons between a variety of countries, this bias should not occur. Respondents will, in this case, compare a foreign country with their home country or with an even larger set of countries. They will be forced to apply the same definition of corruption and make use of the same ethical standard for all countries, which produces valid comparative assessments. However, in this context a second type of bias might arise, originating from the potential dominance of a particular cultural heritage in the sample questioned or because expatriates lack a proper understanding of a country's culture. If this happens, comparative assessments might reflect disproportionately the perceptions of a particular culture, while cultures may differ in their perceptions of precisely where the dividing line between corruption and legitimate and approved social interaction with officials may lie. Such culturally based perceptions would only help our understanding of real levels of corruption to a limited degree.³ While samples which are dominated by a particular cultural heritage are susceptible to this kind of bias, surveys which question local residents clearly avoid this kind of bias.

The strength of the CPI rests with the idea that we include surveys which are not susceptible to the first type of bias. Particularly these are CEER, EIU, PRS, FH and PERC. Since the data provided by these sources refer to assessments by expatriates, they are subject to a

³ Less sophisticated viewpoints towards the CPI alleged in the past that it was driven by the viewpoints of western oriented businesspeople. This viewpoint is certainly wrong.

homogeneous definition of corruption and a consistent ethical standard. The CPI also incorporates the data from the ICVS, IMD, WEF, GI and WB/UB. Since these refer to assessments made by local residents, they are not likely to represent the perception of a certain cultural heritage. The second type of bias can clearly be rejected for these sources.

Since the data from the CEER, EIU, PRS, FH and PERC correlate well with the other data, there seems to be no support for the suggestion that they might be influenced by the second type of bias. Similarly, since the data by the ICVS, IMD, WEF, GI and WB/UB correlate well with the other three sources, the notion that the first type of bias might be present is clearly not supported. The validity of the sources is mutually confirmed and prevalence of the potential biases mentioned before rejected by their high correlation. The approach clearly suggests that the perceptions gathered are a helpful contribution to the understanding of real levels of corruption.⁴

⁴ A method for rejecting both biases at the same time has been developed and tested at an internet site of Göttingen University (www.uni-goettingen.de/~uwwv). Between January 1997 and May 1998 540 replies to an interactive questionnaire were obtained. Users of the internet with an interest in corruption have been asked to approach a page, where the following question has been posed: *"You enter a public office which is authorised to grant licenses and permits (e.g. the license to conduct business). After you waited for a long time you are expected to pay a bribe and are told that otherwise you will not receive the license. According to your perception, in which countries may this (i.e. the asking for bribes by public officials) happen? On the other hand, where do you consider it to be unlikely?"* Three alternatives "often", "sometimes" and "rarely" are given thereafter, which are supposed to be filled with country names.

Each respondent is supposed to assess all countries where he or she obtained first hand experience (the resulting index correlates 0.93 with the 1999 CPI). Thus, the first bias is not likely to occur. The second potential bias can be checked by constructing sub-samples of respondents, dependent on their residence, origin or profession. The resulting index of such a subgroup can be correlated with that of the full sample of respondents, indicating potential differences. Since the resulting correlation for non-western residents and that of people with non-western origin are all higher than 0.95, there are no indications that the cultural heritage is a crucial determinant for the assessment of levels of corruption. The assessments by business people correlate 0.98 with the full sample, indicating no impact of profession on the assessment of corruption.

3.3 The Role of the Media

Another potential problem with the collection of perceptions may arise from the possibility that respondents do not report their personal experiences but rely on media coverage and reports obtained from others. Certainly this influence cannot be excluded and necessarily contributes to perceptions. Yet in its extreme form such an influence may suggest that respondents rely only on hearsay. The potential problem with this influence is that the assessment of a country might then reflect the quality of the press in uncovering scandals, and particularly its freedom to do so. Countries that suppress a free press may escape a bad reputation for corruption among their population. Such an influence would certainly undermine the validity of the CPI and its usefulness as an aid to understanding real levels of corruption.

Investigating whether such an influence might affect the CPI, it is worthwhile to observe that some sources may be more influenced by hearsay than others. It is particularly interesting that the question posed by the ICVS clearly relates to personal experience as opposed to hearsay. The large correlation of ICVS with other sources (commonly higher than 0.8) indicates that hearsay does not appear to be an important influence for the overall CPI.

Another interesting piece of evidence comes from the WEF. In 1999 respondents were asked to state the level of agreement with the following statement: "Personal bribes and kickbacks to senior politicians is rarely alleged in public discussions and rumors." The extent of agreement to this statement clearly reflects an assessment of hearsay rather than personal experience. But the resulting index correlates 0.88 with that provided by WEF which enters into the CPI. This

Since this type of sample design is not a statistically robust approach, this validation step must be seen to be experimental at this stage. This is also the reason for not including these data into the 1999 CPI nor to claim statistical robustness of the results. The results are reported here so as to give a first impression of the necessary methodology and to provide first insights into the suggestive results.

emphasizes that the level of rumors about corruption is not a bad indicator for actual experience about levels of corruption.

4. The Index

4.1 Weighting the Data

With the various sources having some differences with respect to sample and date, a number of ideas have been considered for weighting the sources before aggregating them. One possibility was to weight them according to the number of replies collected by each source. However, this would mean that the GI and ICVS data would overinfluence the results, particularly if seen against the expert assessments conducted by PRS, FH and EIU. If this line were pursued, it would mean that the views of an individual selected at random would have the same quality as an expert assessment made after country-specific analysis and peer review. This approach was therefore not convincing.

Also the idea of assigning a higher weight to more recent data and lower weights to older data was explored. An index which is weighted using such a technique (e.g. taking the weights 3,2,1 for 1999, 98, 97 respectively) correlates 0.998 with an unweighted index, indicating that the differences are negligible.

Another methodology for aggregating governance data has recently been suggested by Kaufmann, Kraay and Zoido-Lobaton [1999], based on a formal model. The authors assume that each source is a noisy indicator for actual levels of corruption, which is the "unobservable component" they seek to determine. Based on this model an average score and a measure of precision is obtained for a large variety of countries.⁵ Those sources which then better correlate

⁵ With respect to measures of precision, some not necessarily realistic assumptions had to be introduced. The measures of precision therefore represent a best-case scenario and are not unbiased. In addition, they neglect the standard deviation between sources, i.e. that precision should be lower in countries where sources differ considerably in their assessment.

with the resulting aggregate index receive more weight, while those which contribute less viably enter into the index with less weight. The quality of the sources is therefore determined endogenously and is not an expert's opinion on a source's validity and reliability. While there may be a point in taking this approach, weighting can be biased if the sources are not independent of each other. It may occur that the sources that are least independent — for example because they use other sources as their benchmark or sample people who have little first-hand experience — are given higher weights than those who engage in discovering original insights.⁶ This weighting system would be in contrast to experts' viewpoints regarding the quality of sources. Given this disadvantage it was decided that this approach should not be adopted for the TI-CPI. In the end, it remains preferable to adopt the simple approach of assigning equal weights to those sources which have been found to meet the criteria of reliability and professionalism. Other procedures may have their merits, but this simple averaging system is easiest to explain to a broad public.

It was suggested in this context that data from various years provided by the same source should not obtain the same weight as other data. One may adhere to the viewpoint that the data provided by an institution is independent to that from another institution, but the same independence may not prevail for surveys originating from the same institution. But this argument may push too far an issue which is in fact difficult to assess. Some institutions may not be fully independent in the assessment of their data; they may lean on the data produced by others in reaching a conclusion. Since the matter of independence is difficult to quantify, there was no clear argument in changing the methodology used so far. As a result of giving each survey an equal weight, some institutions obtain a larger weight than others. While other approaches can certainly

⁶ Apart from these methodological issues, the approach by Kaufmann, Kraay and Zoido-Lobaton [1999] included sources which mixed aspects of xenophobia and nationalism alongside with corruption or which intended to measure changes instead of actual levels of corruption. It was interesting to observe that as a

be justified there is also some rational in this. It reflects previous decisions by the Transparency International Steering Committee that continuous annual surveys are superior for our purposes than one-off surveys: they may have gathered more expertise in providing their service and their inclusion helps to avoid abrupt year-to-year changes in the CPI. In addition to that, surveys may be seen to be superior to expert assessments because the methodology of producing data is more transparent and subject to a clear procedure as opposed to expert viewpoints.⁷

4.2 Standardizing

Since each of the sources uses its own scaling system, aggregation requires a standardization of the data before the mean value for each country can be determined. For all sources not already standardized for the CPIs of previous years, the 1998 CPI was the starting point for this process. It had a mean value of 4.88 and a standard deviation of 2.40. Each of the sources naturally had different means and standard deviations. Yet standardization does not mean that each source is given the same mean and standard deviation, since each source covers a different subset of countries. Instead, the aim of the standardization process is to ensure that inclusion of a source consisting of a certain subset of countries should not change the mean and standard deviation of this subset of countries 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. The aim here is that a country not be "punished" for being compared with a subset of relatively uncorrupt countries, nor

consequence the resulting approach tended to rate poor countries more corrupt than rich countries — a misperception which many people may not be immune to, but which should clearly be avoided.

⁷ Expert assessments may also be less independent than surveys because they tend to be well informed about the results of other sources.

rewarded for being compared with a subset perceived to be corrupt.⁸ In order to achieve this, the mean and standard deviation of this subset of countries must take the same value as the respective subset in the 1998 CPI.

An example can illustrate the standardization. In 1999, IMD assessed France with a value of 5.23 on a scale between 0 and 10. At first, a common subset of countries was determined, countries which belong to both the IMD 1999 and the 1998 CPI. The means and standard deviations in each of these sources was determined. In the IMD 1999, these countries had a mean value of 4.68 and a standard deviation of 2.66, while in the 1998 CPI, these countries had a mean of 6.04 and a standard deviation of 2.47. Standardizing the value for France thus required subtracting 4.68 from the 5.23,⁹ multiplying the result by 2.47, dividing by 2.66 and adding 6.04. The result turns out to be 6.56, the standardized value for France. Applying this to all countries in the subset, the standardized values had a mean of 6.04 and a standard deviation of 2.47, the same values this subset of countries had in the 1998 CPI. The same formula is then applied to all countries covered in the IMD, including those that do not belong to the subset described above. After this is done for all countries and all sources, the index is determined by computing the simple mean for each country.¹⁰

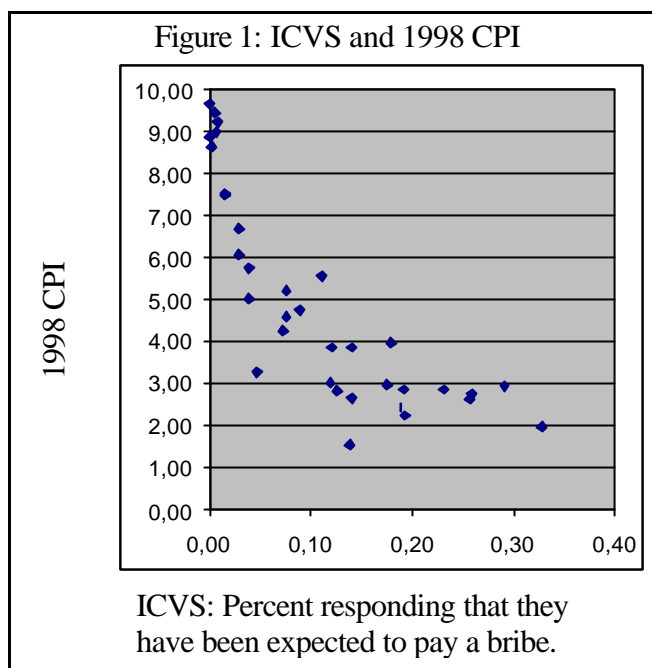
The previous indices relied solely on a technique of standardizing means and standard deviations for the respective subsamples of countries. It was observed in the past that an alternative technique of matching percentiles would bring about largely identical results. Matching percentiles

⁸ It had been alleged in the past that the mere inclusion of countries into some surveys may account for their (low) score. This was called the "curse of inclusion". But by making use of the standardization technique this misrepresentation was avoided.

⁹ With some sources (WB/UB) assigning higher values to more corrupt countries, this value must be multiplied by -1.

¹⁰ A final standardization must be undertaken, since the aggregate may again differ with respect to mean and standard deviation as compared to the previous years index.

is superior in combining indices which are differently distributed. But, as it makes use of ranks and not scores of sources, it loses some of the information inherent in the sources. The general approach is therefore that it is preferable to rely on the described standardization technique, except where the distribution of a source clearly differs from that of the CPI.



One such source with a clearly different distribution is ICVS. As can be depicted from the graphics, the relationship between ICVS and the CPI 1998 does not appear to be linear. There is no other source where this is so clearly the case. It was decided that for ICVS the method of matching percentiles would be applied. For this technique again the common subsamples of the ICVS and the 1998 CPI are determined. Then, the largest value in the 1998 CPI is taken as the standardized value for the country ranking best by ICVS. The second largest value is given to the country ranking second best, etc.

For IMD, CEER and PERC, this standardization procedure did not change the values significantly, since the data was already delivered on a scale between 0 and 10. This contrasts to the values provided by WEF and WB/UB who report the data on a scale between 1 and 7. Likewise PRS and EIU provide assessments ranging between 0 and 6 and between 0 and 4, respectively. The GI data was obtained in raw format and processed with the help of correspondence analysis prior to applying the outlined standardization procedure, see Lambsdorff [1998b]. The original data by Freedom House were not given in numerical format but a "broader alphabetical grade" assigned to the respective categories. This implies that FH does not invite for a

"cardinal" interpretation of their assessments which is required for the normal standardization methodology. By making use of a methodology of matching percentiles only the ordinal information by FH would be used. But also a normal standardization of the data provides an indicator which correlates 0.992 with the one obtained from matching percentiles. Given this it was decided to keep the normal standardization technique.

4.3 Presentation

The 1999 CPI will include all countries for which at least three sources had been available.¹¹ Some scholars had argued in favor of extending the index to include also countries for which less than three sources are available. There are undeniable merits to this. A larger list of countries would further facilitate the usage of the CPI in academic research. There has been an immense research activity based on the CPI, but the limited number of countries was sometimes felt to present a slight disadvantage. Above that, it was observed that limiting the index to countries where sufficient information is available would exclude particularly countries perceived to be corrupt. It was argued that this may mislead the public.

But these arguments must be valued against the respective disadvantages of a further expansion. In public debate, measures of precision are commonly not well taken into account — irrespective of the immense effort TI has put into the presentation of the CPI in the past. Most media still misinterpret the fact that being worst in the index does not mean being most corrupt in the world. In the 1998 CPI this position has been occupied by Cameroon. Expanding the 1998 index to 150 countries would assign this position to another country. There is no clear hint that an expansion of the index would help to avoid the misinterpretation by the media. Even in a larger index there are still more than 50 countries missing because no assessment at all is available, e.g. for Rwanda, Burundi, Djibouti, Central African Republic, Laos and Cambodia. It could easily be

¹¹ For all countries in the CPI at least three sources from three different institutions have been available.

that one of these countries is perceived to be the most corrupt country. Calling a country the worst in the world remains equally inappropriate even for this larger index. But an expansion may even more invite for this interpretation, because an index of 150 countries would include most countries in the world.

The method to avoid this misperception has been to restrict the index to those countries where sufficient information is available. It therefore makes sense to stick to this established guideline and include only those countries for which at least three sources were available. Since those countries left out of the index are on average perceived to be rather corrupt, there emerges an inadequate comparison of a country to the rest of the world — an interpretation which TI did not invite for but which some media was engaged in. It may be worthwhile to note that all 150 countries would on average score 4.0. This figure may serve as a benchmark value. Particularly it illustrates that countries not being included into the CPI should not interpret this as a particular type of qualification or disqualification.

On the web-sites we will include the maximum and minimum (standardized) values for each country. These are intended to better illustrate the margins of error that are associated with a country's score. Moreover, some observers may be interested in the amount of independent institutions that contributed to an average value and not only the total number of sources. This value will also be reported. Apart from that, the CPI will continue to rank countries and assign scores with one digit, as we have done in the past.

References

- HEUNI (1998), "Crime and Criminal Justice in Europe and North America.", *European Institute for Crime Prevention and Control, affiliated with the United Nations (HEUNI)*, P.O. Box 161, FIN-00131 Helsinki. Ed. by K. Kangaspunta, M. Joutsen and N. Ollus. The study can be obtained at: <http://www.vn.fi/om/suomi/heuni/news/ccj.pdf>.
- Kaufmann, D., A. Kraay and P. Zoido- Lobaton (1999), "Aggregating Governance Indicators." World Bank Policy Research Working Paper No. 2195 (Washington D.C.: The World Bank).
- Lambsdorff, J. Graf (1998a), "The Transparency International Corruption Perceptions Index. 1. edition 1995" *Transparency International (TI) Report 1996*, 51-53. "2. edition 1996", *Transparency International (TI) Report 1997*, 61-66. "3. Edition 1997", *Transparency International (TI) Newsletter, September 1997*. "4. Edition", September 1998. A complete documentation can be obtained at: <http://www.uni-goettingen.de/~uwwv>.
- Lambsdorff, J. Graf (1998b), "Corruption in Comparative Perception." *Economics of Corruption*, ed. by A.K. Jain (Dordrecht: Kluwer Academics), 81-109.
- Lancaster, T.D. and T.G.R. Montinola (1997), "Toward a Methodology for Comparative Study of Political Corruption", *Crime, Law and Social Change*, XXVII, 185-206.
- Rose-Ackerman, S. (1999), *Corruption and Government. Causes, Consequences and Reform* (Cambridge: Cambridge University Press).
- United Nations (1999), "Global Report on Crime and Justice." *Report by the United Nations Office for Drug Control and Crime Prevention* (New York: Oxford University Press).

Table 2: Sources for the TI-CPI 1999

Number		Source	Name	Year	Internet Address	Who was surveyed ?	Subject Asked	Number of Replies	Coverage
9	8	Institute for Management Development	World Competitiveness Yearbook	1999	www.imd.ch	Executives in top- and middle-management; domestic and international companies	Improper practices (such as bribing or corruption) in the public sphere	4160	46
	1998			4314					
	1997			2515					
7	6	Freedom House	Nations in Transit	1998	http://www.freedomhouse.org/	Assessment by US academic experts and FH-Staff	Levels of Corruption	Not Applicable	28 Eastern European Countries
				1997				26 Eastern European Countries	
				1997				Not available	
5	4	Wall Street Journal, Central European Economic Review	Annual Survey	1997	http://wsje.com/	Business Analysts in Europe and the United States	Effect of corruption on attractiveness of country as a place to do business	Not available	26 Eastern European Countries
				1997				> 34000 (almost 1000 per country)	
				1997				44	
4	3	Gallup International	50 th Anniversary Survey	1997	www.gallup-international.com	The General Public (internal)	A lot, many, few or no cases of corruption for the following groups of people: politicians, public officials, policeman and judges.	40-50 per country	12 Asian Countries
				1998				280	
				1999				280	
3	2	Political & Economic Risk Consultancy	Asian Intelligence Issue	1997	http://www.asiarisk.com/	Expatriate Business executives	Extent of corruption in a way that detracts from the business environment for foreign companies	40-50 per country	12 Asian Countries
				1998				280	
				1999				280	
1	1	Political & Economic Risk Consultancy	Asian Intelligence Issue	1997	http://www.asiarisk.com/	Expatriate Business executives	Extent of corruption in a way that detracts from the business environment for foreign companies	40-50 per country	12 Asian Countries
				1998				280	
				1999				280	

17	"International Working Group"	1996-1997	16	15	14	13	12	11	10
	International Crime Victim Survey	1996-1997	Economist Intelligence Unit	World Bank/Basel University	Political Risk Services	World Economic Forum			
			Country Risk Service and Country Forecast	World Development Report, Private Sector Survey	International Country Risk Guide	African Competitiveness Report			
			1999	1997	1999	1998	1999	1998	1997
	http://ruljis.leidenuniv.nl/group/fcr/www/icvs/index.htm		www.eiu.com	http://www.unibas.ch/www/wifor/staff/bw/survey/index.html	www.prsgroup.com	www.weforum.org			
	General Public		Expert Staff Assessment (expatriate)	Business Executives	Expert Staff Assessment	Senior business leaders; domestic and international companies ¹²			
	During 1995, has any government official, for instance a customs officer, police officer or inspector in your own country, asked you or expected you to pay a bribe for his service?		Assessment of the pervasiveness of corruption among politicians and civil servants	Irregular, additional payments are common and represent an obstacle to doing business.	Assessment of "Corruption in Government"	Irregular, additional payments connected with import and export permits, business licenses, exchange controls, tax assessments, police protection or loan application.			
	ca. 40,000		Not Applicable	> 3500	Not Applicable	582	3934	3167	2778
	43 ¹⁴		115	74 ¹³	135	20	59	53	56

¹² In 1999 26% of the responses came from business leaders of large domestic companies, 42% from significant exporters, 25% from multinationals and 7% from government officials.

¹³ The World Development Report includes only 69 countries. Further 5 countries have been assessed by Basel University.

¹⁴ Overall 45 countries had been included, but Scotland, Northern Ireland and England & Wales are not counted separately here.