Measuring the Immeasurable:

Boundaries and Functions of (Macro) Corruption Indices

Fredrik Galtung[*]
Tiri

The CPI has been cited in thousands of newspaper articles. It is cited almost on a daily basis. In some contexts, as illustrated above, it can be credited with influencing a country's political trajectory. The CPI was a watershed in the mid-1990s. It contributed towards the formation of a global movement and widespread consensus against corruption. The CPI was a formidable instrument in raising awareness about the international scope and shared burden of corruption and driving corruption onto the front pages of newspapers throughout the developing world. The CPI levelled the playing field by comparing, for the first time, disparate and distinct countries on the same scale. The international shaming that ensued, encouraged a race to the top, i.e. to lower levels of corruption. The race was on the international stage for some (e.g. a senior advisor to South Korea's prime minister sharing his stated goal for Korea to be among the top-15 countries within 5 years[1]); frequently the race was regional (e.g. between Hong Kong and Singapore; between Kenya and Uganda; Hungary and the Czech Republic, etc.); and for some selected countries at the bottom of the league table (e.g. Bangladesh, Nigeria and Paraguay) it has spurred a determination to shed the label of being 'one of the world's most corrupt countries.'

With the advent of the UN Anti-Corruption Convention in 2003, international efforts to curb corruption entered the phase of implementation and enforcement. International efforts moved on from the short-term goals of setting international standards and general awareness raising. The value of the CPI as an awareness raising tool is not called into question.[2] This new international context, however, has called for a reassessment of the functions and limitations of the CPI. The failings of the CPI can be grouped under seven headings. They are referred to hereunder as the CPI's seven failings:

[1] *Interview* with Mr. Choi Byung-Rok in Ottawa, Canada, 14 September 2000.

[2] In 2002 TI was given an Award for Agenda Setting by Media Tenor, an Institute for Media Analysis in large measure reflecting the impact the CPI has had in the world media.

CPI Failing nr 1: Only Punishing the Takers, not the Givers or Abetters

CPI Failing nr 2: Irregular and Uncontrolled Country Coverage

CPI Failing nr 3: Biased Sample: More Than 90% of the World is Missing

CPI Failing nr 4: Imprecise and Sometimes Ignorant Sources

CPI Failing nr 5: Far Too Narrow and Imprecise a Definition of Corruption

CPI Failing nr 6: Does not Measure Trends: Cannot Reward Genuine Reformers!

CPI Failing nr 7: Guilty by Association - Aid Conditionality

The first five failings are significant. Yet a nuanced and reflected understanding of the CPI recognises that no single indicator can fulfil all expectations or counter every criticism. The most effective response to these 'failings' is to recognize the inherent limitations of any given instrument and to balance it, where possible, with other indicators. The last two failings, however, are far more problematic. As the following pages will outline, they contribute to undermining the very rationale for the CPI, which is to be a tool "to *create change* towards a world free of corruption.'[3]

CPI Failing nr 1: Only Punishing the Takers, not the Givers and Abetters

The first criticism of the CPI is that it is one-sided. Mahathir Mohamed, then prime minister of Malaysia, responded critically when informed of the CPI at a press conference: "Who are they to determine how our people should live? This is our country, so we are the ones to decide.'[4] Mahathir suggested that 'a movement to monitor the index of European nations on racial discrimination, corruption and morality be set up in Malaysia,' a 'Europe Watch.'[5]

The major omission is not that the CPI ignores Europe. Many European countries score poorly on the CPI, ranking worse than a number of poorer, less developed countries in Asia, Africa and Latin America. The blind spot is that while CPI casts a spotlight on the major bribe takers of the world, it lets the major bribe givers and safe havens of looted funds off the hook. To correct this imbalance, I was responsible for an effort to develop a brand new index for TI in 1999 to track international bribery. The result was the Bribe Payers Index of leading exporting countries. The imbalanced message of the CPI was to some extent corrected through the commission of this (relatively expensive) new instrument. Yet TI has only produced two BPIs in the ten years between the first launch of the CPI in 1995 and 2004. The BPI remains virtually unknown in the wider public and is generally ignored in the world press.

Furthermore, no instrument has been developed that measures or ranks the willingness by major industrialised and 'clean' countries, like Switzerland, the United Kingdom, USA, but also Singapore, Botswana and Chile, to provide banking and investment safe havens for looted public funds and corruptly obtained fortunes. Numerous reformers in Southeast Asia complain that while Singapore undoubtedly deserves the moniker as Asia's 'cleanest country'

---

[3] From TI's mission statement: http://www.transparency.org/about_ti/mission.html [11.10.04].

[4] 'Don't Tell us how to run our Country,' *The Straits Times* (Malaysia), 5 June 1996.

[5] A. Abdullah and P. Ang, 'PM: Set up Malaysia Watchdog on Europeans,' *New Straits Times* (Malaysia), 5 June 1996.

it has also become a safe haven for corrupt politicians and dealmakers in the region to deposit or invest their looted fortunes. Similar complaints are made in Africa about Botswana and South Africa; and about Chile and the United States in Latin America.

TI was initially formed as the 'coalition to curb corruption in international business transactions.' But has TI done enough to highlight the mechanisms and extent of the involvement by major corporations and banks in industrialised countries in facilitating the infrastructure of international bribery? By producing only two BPIs in the ten years, may TI have unwittingly contributed to a biased assessment that places the overwhelming burden for reform on developing countries (with their comparatively weaker institutions and governance structures), and with far less pressure being exerted on the industrialised world?

There is a more general sense in which the CPI and its attendant sources may be seen as unfairly biased against poor countries. While corruption has been shown to contribute to perpetuate maladministration and underdevelopment there is also ample evidence to support the notion that poverty itself contributes to corruption. If this is the case, is it fair to compare countries with widely differing levels of industrialisation and GDP per capita on the same scale, almost inevitably reinforcing the notion that all good things go together at the upper end of the scale and all bad things come together at the bottom. One solution, inspired by an experiment in British paper, *The Observer*, in collaboration with Amnesty International, would be to weight the corruption scores with a development indicator, like the UNDP's *Human Development Index* (HDI). The logic, according to the authors of this new index, was that 'it would be unfair to compare the impoverished Rwanda with oil-rich Algeria without placing the human rights records of both countries in their economic contexts.'[6] While the approach has a moral appeal, multiplying a corruption score with the HDI produces nonsensical rankings, where countries like Senegal and Gabon become far better placed than advanced industrialised countries. A more practical solution is to produce separate league tables, breaking down the world by geographical regions on the one hand and by levels of development on the other (the HDI divides the world into countries with high, medium, and low levels of development).[7] Comparisons within these categories and intra-regionally are meaningful and they have been used in presentations of the CPI findings.

CPI Failing nr 2: Irregular and Uncontrolled Country Coverage

The CPI requires a minimum of three sources for a country to be included. This is one of the CPI's undoubted strengths. With the exception of data generated in the context of the Bribe Payers Index survey, TI does not commission its own data sources for the CPI. It relies entirely on independently conducted surveys and expert polls. As a result, a significant number of countries cannot be included in the CPI. In 2003 some 133 countries were scored in the CPI. Based on UN membership alone, this meant that 58 countries were missing from the Index.

The failing of irregularity cuts three ways. The first is that countries may be unfairly saddled with the tag as 'world's most corrupt country' or 'the most corrupt country in the Middle East,' when numerous other countries, perhaps even more corrupt, are missing from the ranking.

---

[6] John Sweeney, Peter Beaumont, and Leonard Doyle, 'This is the World Cup that no Country Wants to Win,' *The Observer*, 28 June 1998.

[7] I am grateful to Amartya Sen for his comments on this point.

Second, the reliance on secondary sources means that TI cannot control countries dropping out of the Index if the minimum number of sources is missing. Government tend to respond to this with satisfaction. This occurred in the case of Kenya in 1997. Officials were understood to have briefed foreign diplomatic staff and journalists that year that Kenya's absence from the Index meant that corruption was under control and no longer needed to be monitored.

Third, in countries missing from the CPI, civil society activists, and some foreign aid donors, have found their reform efforts frustrated by governments unencumbered by the pressure and spotlight the CPI generates.[8] Because of the undoubted significance of the CPI as a catalyst for reforms (witness Pakistan, Bolivia, Cameroon and Nigeria in the examples above) it is worth asking whether TI has a responsibility to ensure that the overwhelming majority of countries in the world are included in the CPI.[9] Short of altering its methodology or lowering the entry barrier, the only way of ensuring more extensive and reliable country coverage is for TI to commission regular independent surveys, with an emphasis on those regions that are most poorly covered (significant parts of Africa and the Middle East in particular).

CPI Failing nr 3: Biased Sample: More Than 90% of the World is Missing

The CPI has frequently been criticized for being culturally biased. None too surprisingly, the argument is most commonly made by countries that are the victim of a poor score. This criticism is largely unfounded. According to Saudi Arabia's *Arab News*, for example:

> [Corruption] is a hopelessly subjective concept. What might be considered corruption in Denmark and Sweden or indeed Berlin, might be standard practice in some other countries… [The Western press] talk about bribes and backhanders, when often all that is happening is that commission is being paid for having helped oil a deal […] What is wrong about this particular report is that is [sic] it adopts its own, culturally subjective definition of corruption and then effectively condemns those who do not conform to it. It is an ugly and patronizing attempt to impose moral viewpoints that are the West's alone.'[10]

If the CPI only contained data from national samples, scores could be significantly affected by the level of tolerance for corruption – as well as its visibility, the investigative zeal and integrity of the media. If the CPI only contained the assessments of foreigners, one would be left wondering whether nationals would assess their countries differently and whether nationals of country x can be trusted to provide a fair assessment of country y. Furthermore, since such a foreign sample would more than likely be over-represented by people from the traditional OECD countries, the sample could also be criticised for its 'Western' bias as

---

[8] E.g. in the first effort of its kind, in 2003 the USAID office in Beirut commissioned Middle East survey in the hope of including not only Lebanon but also several other countries from the region. The World Bank, which has also commissioned several international corruption surveys, does not appear to have done so with the same motivation.

[9] See, also, Report on Workshop of 29-30 April 2004, *Good Governance and Behavioural Change*, Phnom Penh, Cambodia: "The CPI is of considerable interest to private corporations – especially when considering whether to invest in a country. Unfortunately, since there is little foreign investment as yet in Cambodia, there have been few corporate donors to pay for such surveys. Nevertheless, it is anticipated that sufficient survey work will have been undertaken to allow Cambodia to be ranked on the CPI within the next 2-3 years."

[10] 'Survey Mania,' *Arab News* (Saudi Arabia), 2 August 1997.

Mahathir did. Mahathir even characterised Malaysians who believe 'such groups [as TI] as people who still have a 'strong' colonial mentality.'[11]

Since the CPI contains both sources based entirely on the assessments by foreigners or expatriates (EIU, FH, PERC, PRS), and sources that are based on samples of nationals (ICVS, IMD, WB/EBRD, WEF) there is no bias in the sample in favour of one or the other view.[12] This is one of its strengths. Furthermore, the divergence of opinion appears to be minimal since the average correlation between the two groups over several years was 0.86. This indicates that on the whole, what counts as corruption in one part of the world, is understood similarly elsewhere. The degree of tolerance and acceptance of corruption may well vary significantly (and the Global Corruption Barometer tends to support this notion), but the characterisation of a country as corrupt is remarkably consistent.

The skewed sample of the CPI is both a strength and its most significant bias. Of the 17 different institutions providing data for the CPI since 1998 only two did not have a private sector bias: Freedom House and Columbia University's State Capacity Survey (CU).[13] Freedom House uses the assessments of in-house experts as well as academics and their findings are not primarily aimed at a business audience. The CU index draws on US-resident policy analysts, academics and journalists. The remaining fifteen institutions either use a sampling frame consisting of business people and/or explicitly target their findings to benefit corporations and institutional investors.

For the robustness of the CPI, the advantages of a homogenous sample are evident. Intuition dictates that business people are far more likely than a random population sample to have first-hand experience and reliable second-hand knowledge of corrupt practices. Their experience, particularly if they have some international experience, will also have the advantage of being comparative. A general respondent with little or no international experience would not be able to compare the level of corruption with any other country and, a priori, be more liable to influence by the media and the political and economic climate in the country.

This homogeneity, while undoubtedly advantageous, also generates a serious sample bias, which is a genuine failing of the CPI. The sample is not only private sector oriented, it is also overwhelmingly male and economically well off. Effectively, this means that this most influential of indices ignores the experiences and perspectives of most women, and of the poor and disenfranchised. It also means that the interests of 'unofficial businesses', which employ the overwhelming majority of the population in poor countries, are ignored.

TI fights corruption in the interest of a five-pronged agenda. At the centre of this agenda is the aim of bringing 'transparency to international trade and commerce.'[14] TI also aims to 'reduce poverty and diminish social injustice' and to 'build democratic and open government.' This analysis suggests that only in a limited sense does this biased sample and

---

[11] 'Don't Tell us how to run our Country,' *The Straits Times* (Malaysia), 5 June 1996.

[12] See Johann Lambsdorff's chapter and the Transparency International website for a glossary of the acronyms for the sources. The sampling frames of ICVS, IMD, WB/EBRD, and WEF consist exclusively of residents living within the countries beings surveyed. Some respondents are likely to be foreigners, however, and in the case of IMD, WB/EBRD, and WEF, which consist of samples of high to mid-level business people,

[13] In previous years, the CPI has included the findings of an international survey by Gallup International, using representative samples of the general population.

[14] http://www.transparency.org/about_ti/mission.html#values [10.10.04]. The other two aims are to contribute to sustainable development and clean environment; and to strengthen global security.

questioning advance the interest of bringing transparency to international trade and commerce. It does nothing to deepen an understanding of corruption that might advance the other agendas set for itself by TI.

CPI Failing nr 4: Imprecise and Sometimes Ignorant Sources

The high level of variance – the lack of consensus - between the sources has been criticized. According to Kaufmann, 'many experts suggest using caution when interpreting TI data and point out that composite corruption ratings are inherently imprecise.'[15] *The Economist* wrote that 'a combination of few sources and high variance between them makes it hard to read much into simple averages, and hence into the rankings.'[16] Kaufmann et al. contend that the data on corruption is only good enough to divide countries into three groups: the 20 or so least corrupt, the 20 or so most corrupt, and the vast majority in between.[17] This criticism of the CPI is only partially valid.

Grouping countries into broad categories would not avoid arbitrariness. Whereas zero and one hundred degrees Celsius are measures with a certain utility related to boiling and freezing points for water, a ten-point scale of corruption has little intrinsic value. Placing a country in the highly corrupt category because it scores 2.5, while placing another country in the less corrupt grouping because of its score of 4 are not meaningful categorisations per se. They are merely relative scores. Moreover, a minor change in a country score could mean that a country leapt from one category to the next, generating a media reaction that something highly significant had taken place. The underlying social and economic reality of the country is unlikely to have changed to any significant degree. One could in theory use confidence ranges to indicate, e.g. with an eighty or ninety percent confidence range whether a country should be characterised as 'honest', 'relatively honest,' 'relatively corrupt', or 'corrupt,' or by signalling the range using a traffic light as an illustration.[18] Figure 2 of Lambsdorff's chapter illustrates these confidence intervals for the CPI.

The criticism contained in the pages of *The Economist* would have more validity were it not an argument one could make with equal validity about a great number of other, considerably more influential aggregate statistics, such as national accounting figures. In India, where around a third of the world's poor live, there has, in the course of a number of a years, emerged a growing discrepancy between the national accounts and household survey data.[19] Such accounting discrepancies are also to be found in advanced industrialised countries. Canadian estimates of the unrecorded (but not necessarily illegal) economy range from 0.6 to close to 22 percent of GDP.[20] Studies for the US put the figures at between 4 and 33 percent, Germany's between 3 to 28 percent and the UK's at 2 to 15 percent. Among industrialised countries, Italy is widely viewed as having the largest unrecorded economy, perhaps as much as one third of its recorded GDP. Among industrialised countries, Italy is

---

[15] Daniel Kaufmann, Aart Kraay, and Pablo Zoido-Lobatón. 1999. 'Aggregating Governance Indicators.' *World Bank Policy Research Working Paper.* [paper updated and revised in 2002]

[16] 'A Guide to Graft,' *The Economist,* 30 October 1999.

[17] Kaufmann op. cit.

[18] Suggestion made by Daniel Kaufmann in a seminar on 20 February 1999, Washington, DC.

[19] Angus Deaton. 2000. 'Counting the World's Poor: Problems and Possible Solutions.' Research Program in Development Studies, Princeton University, *Unpublished Monograph.*

[20] Rolf Mirus and Roger S. Smith, 'Canada's Underground Economy: Measurement and Implications,' *The Fraser Institute*, 2002.

followed by Spain, Belgium, and Sweden, whereas Japan and Switzerland (with ca. 4 percent) are viewed as among the economies with the smallest unrecorded economies.

Italy and the US are the only industrialised countries to officially adjust their GDP figures to allow for the shadow economy.[21] When Italy did so in the late 1980s, it was characterised as 'il sorpasso'[22] since the country overtook the UK in the size of its GDP. The measurement of GDP statistics for poor countries, which are presented without qualifications for variance, shadow or unrecorded economies, e.g. in International Monetary Fund or World Bank annual reports, are all the more unreliable.[23] What is true for GDP statistics is all the more so for trade and investment data.[24]

The CPI is not particularly precise. The CPI has the advantage, however, of being transparent and honest about this failing.[25] When it comes to measuring a phenomenon as contested as corruption, especially on an aggregate level, rather than undermining the index, such openness may even bolster its credibility. The CPI does not claim to be more precise than it actually is. In the context of the measurement of corruption a degree of variance is evidently acceptable. Furthermore, the variance within and between sources contains valuable information. A change in the level of consensus in a country estimate can be a useful barometer for an evolution in the public understanding of corruption for a given country. Rather than being viewed as a vice, this openness is one that other accounting measures, especially at the aggregate level, could perhaps employ with more forthrightness as well.

Moreover, the relative ranking of Singapore, Chile and Botswana, for example, as the least corrupt countries in their respective regions, come as little surprise to most analysts knowledgeable with these regions. Findings that rank Italy as worse than Estonia, Greece worse as than Jordan and Bangladesh as considerably worse than Russia do call into question some widely held assumptions about the relative degree of corruption in different countries. It is these perspectives, based on the perceptions of thousands of business people around the world that contribute to the impact the CPI generates on a public awareness and policy level. The lack of precision goes with the territory. It is a characteristic shared with many other social and economic indicators.

Far more problematic is a basic question of source validity. Symptomatically, there is genuine puzzlement about the CPI scores in a growing number of countries. Knowledgeable people, who travel widely, frequently find it difficult to reconcile the scores

---

[21] See, e.g. http://countrydata.bvdep.com/EIU/Help/measuringeconomicactivity.htm [10.10.04].

[22] 'The takeover', as in a car bypassing another one.

[23] Adriaan M. Bloem, and Manik L. Shrestha. 2000. 'Exhaustive Measures of GDP and the Unrecorded Economy.' *IMF Working Paper.*

[24] A man's shirt, e.g. is recorded as having been manufactured in a given country for trade purposes according to the location where the collar and cuffs were sewn. If the Chinese export quota is filled, an almost finished shirt can therefore be exported to Hong Kong, where a few stitches to the cuffs and collar will be added, for it then to return to China, where it will be finished, packaged and from where it will be transported.

[25] The relative degree of precision of country scores is captured in the CPI press release league table through two additional columns. The first, on the standard deviation, indicates the difference in the values of the sources. The second, the high-low range, provides the highest and the lowest value of the different sources on a standardised basis.

in the CPI with their own experiences. A letter to a business publication in Ethiopia raises these issues[26]:

> Dear Sir,
>
> I read with some puzzlement your recent article titled 'Combating corruption in Ethiopia.' I would like to believe your sweeping statement that institutional corruption is virtually non existent in Ethiopia, but what perplexes me is why in this case the position of Ethiopia in 'The 2003 Transparency International Corruption Perceptions Index' stands at 92 […] It's not only that Ethiopia according to this survey is far below the median which on its own should be a source of great concern but it's a fact too that its current standing is considerably worse than it was some few years back, a development that is very alarming. […] Anyway assuming your (sic.) are sincere in your recent claim, I wonder, do the foreigners behind this international survey have access to particulars that you for some reason do not have, or is it just that they were too lazy to do their homework and instead they relied on partial, erroneous data or maybe just conjured up unfounded statistics? […] If you firmly stand by your claim and find the position of Ethiopia in the above mentioned list far off the mark, the one inescapable question will be what as Ethiopians do we need to do to rectify this misrepresentation that is probably causing us the loss of substantial amount of potential foreign investments?
>
> Kind regards, Yonathan Berhane

Berhane's letter raises the serious issue of source validity. This issue is exemplified by a dataset that was used in the CPI until 2001. Political Risk Services' (PRS) *International Country Risk Guide* has also been widely used as the primary dependent variable for a number of empirical corruption studies. It is also used in Kaufman et al.'s governance indicators. The PRS guide has two main attractions: it covers a wide range of ca. 140 countries and it allows for a time series analysis as the monthly updates of the dataset have been published regularly since 1980. The PRS customer base is essentially corporate, including eighty percent of the Fortune 500 US corporations. Their proprietary data is also made available to researchers at discounted rates.[27] The ICRG composite ratings and forecasts are based on the following formula:

CRR (country x) = 0.5 (PR+FR+ER), where

CRR = composite risk rating
PR = political risk rating
FR = financial risk rating
ER = economic risk rating

The political risk rating, in turn, is divided into twelve sub-components, one of which is corruption. This, however, is only one of several political risk components that contain a question about corruption. The others are Military in Politics, Law and Order, Democratic Accountability, and Bureaucratic Quality. These components correlate well with each other but they are not exclusively concerned with corruption. They cannot, therefore, be used in an analysis of corruption. It is the corruption component alone that is of interest here. Although the CRR and the ICRG are particularly concerned with those risks that 'can make it difficult to conduct business effectively,' the corruption measure is more concerned with 'actual or potential' political and administrative corruption in the form of nepotism, patronage, secret party funding, etc.[28] According to the ICRG *Brief Guide to the Ratings System*, the risk posed by such forms of corruption is that it may provoke a popular backlash or

---

[26] Letter to the Editor of *Capital Ethiopia* of 31 August 2004 (http://www.capitalethiopia.com/letter%20 to%20the%20editorial.htm [11.10.04]). Thanks to Rupert Bladon for pointing out this letter.

[27] www.prsgroup.com [10.10.04].

[28] Ibid.

result in the fall or overthrow of a government, which in turn would have wider repercussions on the investment climate in a country. The corruption risk is first determined by calculating

> how long a government has been in power continuously. In the case of a one-party state or non-elected government, corruption, in the form of patronage and nepotism, is an essential prerequisite and it is therefore corrupt, to a greater or lesser degree, from its inception. In the case of democratic government, it has been our experience, almost without exception, that things begin to go wrong after an elected government has been in office for more than two consecutive terms, that is, eight to ten years [...] The lowest ratings are usually given to one-party states and autarchies.[29]

As a working hypothesis for determining actual corruption levels, this is at best indicative and tentative, and in many instances it is likely to be flawed. There is robust evidence, for example, that the scale of both bureaucratic and top-level corruption in most post-Communist countries in Eastern Europe and the former Soviet Union increased markedly in the first years after 1989.

The risk points per country range from 1 to 6, where the low score represents a very high risk, whereas a high score means the risk posed by corruption is minimal. On this scale, Canada, Denmark, Iceland, the Netherlands, and Sweden are among the handful of countries to obtain a score of 6. In 2002, the Dem. Rep. of Congo, Indonesia, Gabon, Somalia, Yugoslavia and Zimbabwe are among the countries at the other end of the scale. So far, these ratings are not particularly controversial. What is more disconcerting is, e.g. that Botswana, widely considered to be the least corrupt country in Sub-Saharan Africa, obtains a 3, when Congo-Brazzaville gets a 4; Liberia, which barely has a functioning state, even scores 5. Botswana has been ruled by a single party since independence, but corruption levels there are trivial compared to Congo-Brazzaville or Liberia.[30] Norway, in turn, obtains a 5, putting it on par with Liberia, Germany gets 4 and France even scores 3, level with Malawi and Yemen.[31] The editor of the ICRG explains these apparent anomalies:

> It frequently happens under our system that a country with a very low level of 'institutionalised' corruption presents a higher risk in our Corruption component than a country where corruption is institutionalised, the reason being that the political risk posed is greater in a democratically accountable system than in one where democratic accountability is severely curtailed or non-existent.[32]

On the one hand, there is the doubtful validity of using one-party rule, autarchy or indeed the length of a government in office as a proxy for corruption levels. On the other hand, the ICRG corruption component measures the risk posed by corruption to the private sector, not the incidence or scale of corruption per se. Whether or not the researchers using the ICRG in their regression analyses have been aware of these characteristics, they must cast doubt upon the validity of the ensuing research findings. A number of researchers have used a risk indicator that is at wide variance with numerous other measures of the scale or incidence of corruption.

---

[29] Thomas Sealy. 2001. 'International Country Risk Guide (February).' *Political Risk Services* 21.

[30] Botswana is a democracy whose government has not changed since regaining Independence in 1966 and it scores very favourably both on the CPI and in the opinions of those familiar with the region.

[31] In TI's 2003 CPI press release (http://www.transparency.org/pressreleases_archive/2003/2003.10.07 .cpi.en.html) Norway, France and Germany are among the countries cited as having improved their scores in recent years. A significant percentage of the upward trend of the scores for these and other countries would be accounted for by the exclusion of PRS from the data sources.

[32] Thomas Sealy, editor of the ICRG, *email* to the author of 18 May 2001.

PRS is not the only source that raises questions of reliability. The potential problem of circularity emerges with the Economist Intelligence Unit a source that has been a staple of the CPI since its early days. The EIU obtains a 'cleansed' version of the CPI from TI that does not contain the EIU's original data. The EIU thereupon uses this as a benchmark for their own ratings. The problem is compounded by the uneven level of expertise within the EIU. For some countries it is excellent. In other cases, 'some editors don't speak or read the language of the countries they work on and they may have little or no in-country experience.'[33] For the EIU the corruption score is only one of several factors that contribute to a country risk rating. For the CPI, the corruption score is the only one that matters.

CPI Failing nr 5: Far Too Narrow and Imprecise a Definition of Corruption

The sources used in the CPI only make weak distinctions distinguish between different branches of the civil service. If they were limited to any one aspect, e.g. the judiciary or to corruption in political party financing, it would be difficult to justify their inclusion. IMD inquires about the 'public sphere,' for EIU it is the 'public office,' and PRS uses 'government.'

As was described in failing nr 3, the bias of the sources favours corruption as it affects the private sector. The private sector is likely to be particularly concerned about the forms of corruption that affect it most. The World Economic Forum's survey on global competitiveness is particularly detailed, covering 'export permits, business licences, exchange controls, tax assessments, police protection or loan applications.' In other words, any form of corruption that would be likely to be encountered by business people and investors. The surveys may also be concerned about high-level government corruption if it poses a risk to their business interests or if it is a factor in political volatility. They are likely to be less concerned by these factors, e.g. in Indonesia during the heyday of the Suharto years or in China during the early 2000s, when the political environment appears stable and favourable to foreign investors.

Table 1. Illustrative Questions from CPI Sources[34]

Information International	*How common are bribes, how costly are they for doing business, how frequently are public contracts awarded to friends and relatives?*
World Business Environment Survey	*Frequency of bribing and corruption as constraint to business.*
World Markets Research Centre	*Red tape and the likelihood of encountering corrupt officials. Distinguishes between small-scale bribes, larger kickbacks and corporate fraud.*
Gallup International for TI	*How significant of an obstacle are the costs associated with [bribes] for doing business?*

This standardisation has advantages for the CPI. Yet the understanding of corruption, particularly in the context of poorer countries, is inadequately advanced. In order to

---

[33] *Interview* an EIU country editor, London, 8/6/98.

[34] See http://www.transparency.org/pressreleases_archive/2003/2003.10.07.cpi.en.html [12.10.04] for a complete list of the questions.

maximise response rates and in order to serve the interests of their primary clients, the questions on corruption place the respondent in the position of being an unwitting victim of the state (see Table 1 for illustrations). Collusion or agency by the respondent is never explored. Corruption is thereby couched entirely in terms of public extortion. The Bribe Payers Index shares this slanted approach. It asks respondents to rate the behaviour of competitors in areas of business with which the respondent is most familiar. Prosecutors and journalists (indeed, business people as well) know that this only captures one slice of the phenomenology of corruption. China, Panama, Sri Lanka and Syria all shared the same CPI score in 2003 of 3.4. Only one of these countries is a major recipient of Foreign Direct Investment. What is it that makes corruption more than tolerable in one context and yet intolerable in others? The CPI blurs the differences between corruption as an annoyance and corruption from which the private sector genuinely recoils.[35]

Cross-country comparisons of corruption levels in the different spheres of government would be particularly important in the context of reform. They may help address the question of why some venal business environments are more attractive or risky than others. From the standpoint of reformers working to promote a broader public interest, it would contribute significantly towards understanding the real pressure points and sequencing for sustained reform. The CPI in its present form, masking the trees for the forest, does not help to answer these questions.

The CPI would more accurately be called the 'bribe takers perception index' or, even more accurately, the 'extortion perceptions index.' The nomenclature of corruption is much broader than bribery or extortion. Some of the sources used in the CPI make allowances for distinctions between large and petty bribes. None of them distinguish between a wider catalogue of corrupt acts, including nepotism, extortion, patronage, facilitation payments, collusive networks, administrative and political corruption, etc.

The CPI benefits from its homogeneity. The range of activities it describes is narrower in scope and more internally consistent than it would be if it were required to subsume the entire catalogue of corrupt acts in one index. The unintended consequence has been a definition of anti-corruption globally almost entirely in terms of anti-bribery. More specifically, much anti-corruption work has been identified with the reduction of bribes that victimise private investors, investors who are frequently foreign.

Failing nr 1 described the one-sided nature of the CPI. It depicted the way in which this instrument has focused attention on bribe takers while letting bribe givers and abetters have an easy ride. Failing nr 3 expanded on the private sector bias of the CPI, particularly in terms of sampling. Failing nr 5 makes it clear that when the analysis is also limited to bribery (and more accurately to extortion by public officials) the analysis and the ensuing onus for reform have been severely constrained. Other systems of corruption, such as political patronage networks, nepotism or state capture by major private interests have only been analysed in isolation or on a regional basis by other organisations, including some national chapters of TI. Reducing corruption that victimises foreign investors is a significant agenda

---

[35] Illustrative of the fact that leading rankings is an insufficient pre-condition for foreign investments, Haiti topped the International Monetary Fund's 'Index for Trade Openness' in the early 1990s for being the 'best student of the rules of globalization […] Yet over the 1990s, Haiti's economy contracted […] No surprise – if you are a corrupt and misgoverned nation with a closed economy, becoming a corrupt and misgoverned nation with an open economy is not going to solve your problems.' (Tina Rosenberg, 'The Free-Trade Fix,' *The New York Times Magazine*, 18 August 2002.)

for reform. The CPI has contributed to making this the dominant paradigm for reform to the exclusion of other agendas.

CPI Failing nr 6: Does not Measure Trends: Cannot Reward Genuine Reformers!

The CPI's principal flaw is that it is a defective and misleading benchmark of trends. Initially set up to encourage reform, the CPI cannot answer the basic question: 'After four years, are these reforms making any difference?'

Because of the failing of irregularity (nr 2) TI is unable to control the countries that drop in and out of the CPI. If five countries drop out one year and six new countries come in, this will affect the rank order of some countries. To be fair, in its press releases TI warns against misinterpreting such arbitrary changes in the rank order of countries. Despite these warnings, media headlines frequently refer to changes in a country's rank order and the various caveats on TI's website remain largely unreported and widely misunderstood.

Another concern voiced both by reformers in governments and civil society, is that anti-corruption drives may bring corruption into the open and thereby tarnish a country's image, precisely during a period when genuine efforts at reform are introduced. There is no evidence to support this perspective in the history of the CPI. Moreover, the judicious choice of averaging survey data from three years reduces the blips that might be caused by major one-off scandals and investigations. The working hypothesis has also been that significant structural reforms that go beyond targeted and punctual reforms in specific public procurement do take at least a few years to bear fruit.

More troubling, is the fact that changes in country scores cannot necessarily be attributed to independently verifiable changes in the corruption environment. In TI's own words,

> Year-to-year changes in a country's score result not only from a changing perception of a country's performance but also from a changing sample and methodology. Some sources are not updated and must be dropped, while new, reliable sources are added. With differing respondents and slightly differing methodologies, a change in a country's score may also relate to the fact that different viewpoints have been collected and different questions been asked. The index primarily provides an annual snapshot of the views of business people and country analysts, with less of a focus on year-to-year trends.[36]

Trends cannot be assessed in terms of changes in the rank ordering. Nor, however, can they be identified in terms of changes in a particular country's scores. There are two principal reasons why country scores do not provide a cogent benchmark for trends. The first, as outlined above, is that the composition of the sources and methodologies employed frequently change from year to year. There have been numerous changes over the decade since the CPI was first introduced, both in the sources selection and the standardisation techniques used to aggregate the data. There have been three significant changes over this period. The second reason the scores do not provide a useful benchmark for trends is inextricably linked to the standardisation technique. The passage describing the current standardisation technique, matching percentiles, is worth citing in full:

> Each of the sources uses its own scaling system, requiring that the data be standardized before each country's mean value can be determined. This standardization is carried out in two steps. For step 1

---

[36] http://www.transparency.org/pressreleases_archive/2003/2003.10.07.cpi.en.html [12.10.04].

each source is standardized using matching percentiles. The ranks (and not the scores) of countries is the only information processed from our sources. For this technique the common sub-samples of a new source and the previous year's CPI are determined. Then, the largest value in the CPI is taken as the standardized value for the country ranked best by the new source. The second largest value is given to the country ranked second best, etc. 2 Imagine that a new source ranks only four countries: UK is best, followed by Singapore, Venezuela and Argentina respectively. In the 2002 CPI these countries obtained the scores 8.7, 9.3, 2.5 and 2.8. Matching percentiles would now assign UK the best score of 9.3, Singapore 8.7, Venezuela 2.8 and Argentina 2.5.

Matching percentiles is superior in combining indices that have different distributions. But, as it makes use of the ranks, and not the scores of sources, this method loses some of the information inherent in the sources. What tips the balance in favor of this techniques is its capacity to keep all reported values within the bounds from 0 to 10: All countries in the CPI obtain scores between 0 (very corrupt) and 10 (highly clean). This characteristic is not obtained by an alternative technique that standardizes the mean and standard deviation of the subsamples. Matching percentiles, on the other hand, guarantees that all standardized values are within these bounds. This results because any standardized value is taken from the previous year's CPI, which by definition is restricted to the aforementioned range.[37]

Matching percentiles is advantageous in terms of data presentation. The scores will always fall in a range between 0 and 10. With an earlier standardisation technique based on country scores – not ranks – some countries, like Finland, obtained standardised scores of more than 10.[38] At the other end of the scale, Bangladesh emerged with a score of less than zero. While this was statistically sound, in presentational terms it was incomprehensible. Matching percentiles suffers from another defect, however.

When countries are ranked (or indeed scored) by respondents and even expert panels, the natural reflection is to compare a country with its major regional neighbours to have a baseline of comparison. Even if a country implements significant reforms to curb corruption and is assessed as being on a positive trend by some of the CPI's sources, its rank order vis a vis neighbouring countries may take years to change, particularly if other countries in a given region are undergoing reforms as well. In the short term, small improvements in a country's rank order might even be 'punished' by the second phase of the standardisation technique:

Without a second adjustment there would be a trend towards a continuously smaller diversity of scores. If, e.g., Finland were to repeat its score from the previous year, it would have to score best in all sources listing this country. If it scores second to best in any source, the standardized value it obtains after using matching percentiles and aggregation would be lower than its current score. Thus, given some heterogeneity among sources, it seems inevitable that Finland's score would deteriorate. The opposite would be true of Bangladesh, which would obtain a better score if it is not consistently rated worst by all its sources. A second standardization is required in order to avoid a continuous trend to less diversity among scores. However, applying a simple mean and standard deviation technique might again bring about values that are beyond our range from 0 to 10. A more complicated standardization is required for the second step: A beta-transformation. The idea behind this monotonous [*should say* 'monotonic'] transformation is to increase the standard deviation to its desired value, but to keep all values within the range from 0 to 10. [39]

Bangladesh currently has the reputation of being one of the most corrupt countries in Asia. Indeed, it is ranked last in the CPI. This position may be justified. Because of the failing of

---

[37] Johann Lambsdorff, 'Framework Document 2003: Background Paper to the 2003 Corruption Perceptions Index,' Transparency International, Berlin: 7.

[38] In the final press release, the score was capped at 10.0 (see http://www.transparency.org/cpi/2000 /cpi2000.html#cpi [10.10.04]).

[39] Johann Lambsdorff, 'Framework Document 2003: Background Paper to the 2003 Corruption Perceptions Index,' Transparency International, Berlin: 7.

irregularity, one cannot know whether the inclusion of other countries from the region and beyond would moderate the country's standing. Bangladesh is also one of the poorest countries in Asia. With its sizeable population, an active aid community, and all-too-frequent natural disasters it is often in the regional media spotlight, Bangladesh is frequently associated with bad news. Moreover, most other countries in the sub-region are undertaking anti-corruption reforms; and Nepal is not included in the CPI because it only has two sources.

In theory, even if Bangladesh were to make significant improvements in curbing corruption, its score (not only its relative rank), may remain constant for years. Bangladesh would need to improve its image and credibly report to the international business community and financial press that it was making strides against corruption. Institutional reforms alone (new anti-corruption commissions, higher rates of conviction, civil service reform, etc.) would not be sufficient.

For Bangladesh's CPI score to improve something else would be needed: other countries in the region and beyond would actually need to deteriorate or experience significant setbacks. This bad news, also independently verified, would in turn need to be widely disseminated and sustained over a long period (more than 3 years). In other words, for Bangladesh's CPI score to improve, the country would not primarily be competing to improve its own standing, it would also be in a negative race against its neighbours where an overall improvement in the region would be to Bangladesh's disadvantage. (By contrast, if the other countries deteriorate and Bangladesh remain as it is, the score for Bangladesh would rise without Bangladesh having done anything more than stand still.)

By publishing the CPI on an annual basis, TI wittingly sustains this negative race to the bottom. TI accurately acknowledges that the CPI is no more than an 'annual snapshot.' However, publishing the CPI every year it invites erroneous comparisons based on both rankings and scores. The CPI cannot reward reformers because the standardisation technique emphasizes rank ordering over internal reforms. It thereby reinforces negative perceptions about certain countries (and the positive images of other, already well-to-do countries). The CPI, in other words, has become a stick without a carrot. It is all but impossible for countries to improve their scores in the CPI through government reforms and sustained anti-corruption efforts.

Short of ceasing to publish the CPI altogether, a possible way forward for TI, would be for the CPI to be released at most every four to five years. The Bribe Payers Index and a new 'Index of Corruption Safe Havens,' both more robust indices over which TI would have control over data collection and analysis, should be published annually. A variety of other instruments, such as the Global Corruption Barometer, could address the other lacunae of the CPI. When publishing the CPI, TI should make every effort to maximise the country coverage by commissioning its own independent surveys. A modest effort could increase the coverage by at least 35 countries.

In order to put an end to erroneous and misleading comparisons based on country scores, the CPI should simply cease to publish them. Countries should be listed only in rank order. Some misrepresentation would still occur, especially as and when a number of new countries

were added to the Index. But these misunderstandings would be far less severe than the current misinterpretations of country scores.[40]

CPI Failing Nr 7: Guilty by Association - Aid Conditionality

The failing of guilt by association distinguishes between a relatively minor failing and a major problem. The minor one is the CPI's association with numerous correlation and regression analyses produced by economists and political scientists since the mid-1990s. In fairness, a majority of these studies, especially the ones published by economists, the World Bank and IMF, do not use the CPI. World Bank researchers have been able to use their own data sources generated in recent years. A number of studies relying on secondary sources use PRS's ICRG because the wide country coverage and long-standing data set appears to lend itself particularly well to time series analysis.

A reduction in corruption levels by a single point on a ten-point scale is estimated to increase annual GDP growth per capita from 0.6 to 1.8 percentage points[41]; a similar reduction would bring the Gini coefficient for income inequality from 0.9 to 2.1 Gini points[42]; it would reduce the child mortality rate by 1.1 to 2.7 deaths per thousand live births.[43] On a lighter note, a study by the US Federal Reserve Bank St Louis, finds that a belief in hell is negatively correlated with levels of corruption.[44] Numerous studies are based on simple correlations that leave the question of causality open. Others, using more sophisticated techniques, seek to isolate and understand the causalities that lower corruption levels would produce.

The heuristic and political function of these studies has been considerable. These studies provided the prima facie case required, for example, by the Bretton Woods Institutions to legitimise their commitment to anti-corruption. This was a topic they had hitherto explicitly avoided as being 'political' and beyond their remit for fifty years of their existence. A number of these studies share one or more of four flaws. First, if the CPI is the corruption variable, then it cannot be used for time series analyses. Second, if the corruption variable is

---

[40] Since the late 1990s, more detailed in-country diagnostic surveys have been used as one way of measuring improvements and providing more detailed information of use to policy makers. Close to one hundred such surveys have been supported by a range of institutions, in particular the World Bank, but also numerous NGOs around the world. The heterogeneity of approaches and samples means that there is little common basis for comparison between them. And only a few surveys have been repeated with any regularity. Moreover, the CPI continues to dominate the headlines and to be the international media's benchmark by default. The incorporation of the World Bank's 'control of corruption index', in the context of the US Millennium Challenge Account (see next section) may change this balance slightly in the coming years.

[41] Estimates in different studies vary quite considerably, but the relation between reducing corruption and increasing GDP per capita growth is generally seen to be positive (see George Abed, and Hamid Davoodi. 2000. 'Corruption, Structural Reforms, and Economic Performance in the Transition Economies.' *IMF Working Paper* 00/132; C. Leite, and J. Weidmann. 1999. 'Does Mother Nature Corrupt? Natural Resources, Corruption and Economic Growth.' *International Monetary Fund Working Paper* 99/85; Paulo Mauro. 1996. 'Corruption and the Composition of Government Expenditure.' *Journal of Public Economics* 69: 263-279; Vito Tanzi, and Hamid Davoodi. 1998. 'Corruption, Public Investment and Growth.' in *The Welfare State, Public Investment and Growth*, edited by T. Shibata and T. Ihori. Tokyo: Springer Verlag).

[42] Sanjeev Gupta, Hamid Davoodi, and Rosa Alonso-Terme. 1998. 'Does Corruption Affect Income Inequality and Poverty?' *IMF Working Paper* 98/76.

[43] Sanjeev Gupta, Hamid Davoodi, and Erwin Tiongson. 2000. 'Corruption and the Provision of Health Care and Educational Services.' in *The Economics of Corruption*, edited by A.K. Jain. Dordrecht: Kluwer Academics.

[44] 'Belief in Hell Boosts Economic Growth, Fed Says,' *Reuters*, 27 July 2004.

PRS's ICRG, then the study does not address changes in corruption levels; the PRS guide measures risk to investors. Third, the CPI and other aggregate indices, like Kaufmann's governance indicators, are subject to significant variance and standard deviation. They do not lend themselves well to making minute inferences, particularly when the independent variable (e.g. GDP per capita growth) is so inaccurate for many countries (see failing nr 4 above).

Fourth, a large number of countries have remained poor over the last four decades. They tend to share a basket of 'bads' that include weak public institutions, low levels of literacy, poor health statistics, weak judiciaries, etc. Many of these countries are also rated as highly corrupt. Only a handful of countries have made the transition to higher standards of living in the past four decades. Some of these countries have also been rated as highly corrupt. A larger number of countries have enjoyed higher standards of living for more than forty yeas. These countries tend to share basket of 'goods' that is much the inverse of the first category of countries. They are also perceived as largely 'clean.'[45] By the sheer weight of numbers, therefore, regressions, causalities and correlations reveal little that is new.

The CPI's guilt in this case is admittedly a minor one. Where the CPI or PRS's data have been used as the dependent variable the research is faulty. Many of the conclusions are facile and at best heuristic. Much of the research tells policy makers little that they do not already know. As a general rule (and there are some significant exceptions), high levels of corruption are associated with much that is bad. Low levels of corruption tend to be associated with much that is good. The CPI, and the myriad studies it has inspired, provides no actionable outcomes. It can legitimize the case for reform but it cannot genuinely point reformers in any meaningful direction.

The CPI's more serious associative failing is with donor conditionality. In recent years, a number of aid donors have used the CPI in the context of bilateral negotiations. Until recently, such conditionalities were largely informal or politically expedient (as in the case of Kenya during the last years of the Moi regime). With the Bush administration's Millennium Challenge Account (MCA), the use of the CPI in the context of donor conditionality has become a matter of public record. The MCA foresees a fifty percent increase in US core development assistance. While targeted at countries in need, the MCA rewards 'good policy' by rating countries along sixteen 'objective' criteria.[46] One of these criteria is 'control of corruption.' The data source for the corruption scores is generated by the World Bank Institute. The MCA recommends, however, that 'As additional information, the Board may also consider how the country scores on Transparency International's Corruption Perceptions Index.'[47]

The implications of this association may seem benign on the surface. After all, the beneficiaries of the MCA will receive additional US assistance. The candidates who fail, however, are likely to be penalised. Kenya, for example is one of the countries in 2004 that

---

[45] Mushtaq Khan makes a similar argument. Some of his research is on the small group of countries where 'growth has been high despite the presence of substantial corruption' (Mushtaq Khan and Jomo K.S., eds, *Rents, Rent-Seeking and Economic Development: Theory and Evidence in Asia*, Cambridge, Cambridge University Press, 2000: 10).

[46] http://www.mca.gov/ [12.10.04].

[47] Millennium Challenge Corporation, *Report on Countries that are Candidates for Millennium Challenge Account Eligibility in FY 2004 and Countries that are not Candidates Because of Legal Prohibitions: Annex A: Indicator Definitions*, Washington, DC, 2004.

'fails to meet the 'control of corruption' standard required for inclusion in the aid programme. Kenya is one of seven countries likely to be disqualified solely on that basis.'[48]

Leaving aside the other weaknesses of the CPI, failing nr 6 (its inability to reward genuine reform), should disqualify it from playing any role in such decisions. The CPI is a publicly available document and TI cannot, therefore, control the many ways in which the findings may be applied. TI could, however, distance itself publicly from the use of its Index in the context of aid conditionality, an application to which it is wholly unsuited. As was suggested in nr 6 above, TI would do even better by ceasing to publish the country scores.

Conclusion

Clearly, corruption can be measured. The questions remain as to how accurately, and to what effect? The CPI contributed to a shift whereby corruption measurements came to be seen as a primarily descriptive exercise, not a predominantly moral and value-laden notion.[49] Drawing on historical antecedents, there are indications that the use of comparative country rankings was seminal in this shift. One of the major innovations in the public health movement of the 1850s in the United Kingdom was the introduction of the notion of the 'Healthy Districts.' These were 63 registration districts, one tenth of the national total, where the crude death rate was one tenth of the national total. In the ensuing decades periodic reminders were published of the tens of thousands of preventable deaths in various cities around the country that would have been avoided if the sanitary conditions approximated those of the Healthy Districts.[50]

Although a more sophisticated and objective ranking than the CPI, the two indices share a use of comparative measurements across units of analysis (districts in the one case and countries in the other) previously held to have been futile. In both instances, a change in awareness and concrete policy reforms followed.

The CPI played a central role in influencing the global agenda for anti-corruption reform, raising awareness by shattering the taboo and catapulting the issue of corruption to the forefront of national and international discourse. With the advent of the UN Anti-Corruption Convention in 2003, which marked the culmination of many years' efforts to produce an armoury of international and regional anti-corruption norms, anti-corruption has firmly entered a very different phase of enforcement and implementation. The immediate goal is no longer simply to raise awareness about the importance of curbing corruption. Nor is it to establish even more comprehensive and stringent international norms. In this latest phase, one to which many newly elected governments are resolutely committed, the CPI in its present form has become counterproductive.

Collectively, the seven failings of the CPI call for a complete reassessment and an overhaul of this influential social indicator. In particular, the CPI needs to be complemented by other indicators to address vital aspects of the subject that a single index can never hope to capture.

---

[48] 'Kenya Likely to be Disqualified for Share in new US Development Aid,' *Financial Times*, 30 April 2004.

[49] Amartya Sen has described a similar trajectory in the measurement of poverty (1973. 'Issues in the Measurement of Poverty.' *Scandinavian Journal of Economics* 81: 285-307).

[50] Simon Szreter. 1991. 'The GRO and the Public Health Movement in Britain, 1837-1914.' *The Society for the Social History of Medicine*: 435-463.

The sixth failing reveals that the CPI is unsuitable to measure trends or even to capture genuine reforms. It should no longer be published in its present form as it actually undermines the efforts of reformers. Short of abandoning it altogether, the CPI should be published at most every four to five years. The country coverage would need to be global. And the country scores should be removed. Transparency International, independently of its views on the utility of corruption-related conditionality in aid or debt relief to poor countries, should publicly distance itself from, and warn aid agencies against, any use of the CPI in such negotiations. The organisation should similarly brief recipient governments that the CPI is unsuitable in this regard.

The challenge ahead is evident: after ten years it is time to find new measurements!