The *Global Integrity Report: 2009*
Methodology White Paper

The Global Integrity Approach

Global Integrity's Integrity Indicators provide the underpinning for our core reporting and analysis of governance and corruption. The Indicators represent one of the world's most comprehensive data sets providing quantitative data and analysis of anti-corruption mechanisms and government accountability at the national level in diverse countries around the globe. Utilizing our network of more than 900 on-the-ground analysts and our unique scoring system, Global Integrity generates actionable, and action-worthy, data and qualitative analysis drawn from more than 300 indicators to assess a country's anti-corruption safeguards and potential for corruption.

The Integrity Indicators are based on a simple yet powerful concept. Rather than trying to measure actual corruption, considered virtually impossible by experts (e.g., Medard 2001; Sik 2002; Arndt and Oman 2006), Global Integrity quantitatively assesses the *opposite* of corruption, that is, the access that citizens and businesses have to a country's government, their ability to monitor its behavior, and their ability to seek redress and advocate for improved governance. The Integrity Indicators break down that "access" into a number of categories and questions, ranging from inquiries into electoral practices and media freedom to budget transparency and conflicts of interests regulations.

We unpack these concepts by looking not only at what laws or institutions are "on the books" but also by assessing their implementation and enforcement through indicators of staffing, budget support, political independence, and citizen access to the most important anti-corruption mechanisms. Combined with vivid narration prepared by our team of leading journalists in the form of accompanying Reporter's Notebooks, the extensive data provided by the Integrity Indicators informs and empowers citizens, activists, donors, businesses, and governments in each country.

Global Integrity country assessments provide detailed data and reporting on the mechanisms in place to prevent abuses of power and promote public integrity. Using a blend of social science and journalism, in-country teams of independent researchers and journalists report on the *de jure* as well as *de facto* reality of corruption and governance. While the Reporter's Notebook on the culture of corruption often paints a depressing picture (how corruption looks, tastes, feels and smells to the average citizen), the Integrity Indicators identify strengths and weaknesses in the national anti-corruption architecture and serve as a road map for possible reforms. Transparency, both in terms of our methodology and findings, is what characterizes the Global Integrity approach and enhances the robustness and credibility of our findings. As we continue to improve our methodology and learn from our fieldwork experiences, we welcome and value critical feedback.

Research Team Members and Roles

Global Integrity staff in Washington both identify and manage our teams of local, in-country reporters and researchers who are responsible for collecting the raw reporting and data that ultimately generates the Global Integrity country assessments. Headquarters staff recruit, contract, and pay in-country team members; develop and share the methodology with in-country experts; review all data and reporting for completeness and accuracy; manage the peer review process (see below for details); and produce cross-country analysis of the resultant qualitative and quantitative information. An obvious but notable point is that Global Integrity *qua* Global Integrity does not "assign" countries scores in any category, sub-category, or Integrity Indicator. Instead, of all our data and reporting is home-grown, bottom-up information generated and peer reviewed by in-country local experts. Global Integrity staff manage the fieldwork process and provide overarching guidance and logistical support.

Fieldwork and Country Selection

From June to December 2009, Global Integrity conducted field research (assessing the period June 2008 to June 2009) in the following 38 countries or territories:

Algeria, Armenia, Azerbaijan, Bosnia-Herzegovina, Brazil, China, Colombia, Georgia, Ghana, India, Indonesia, Jordan, Kenya, Kosovo, Liberia, Lebanon, Macedonia, Malawi, Mexico, Mongolia, Nepal, Nicaragua, Norway, Qatar, Rwanda, Serbia, Sierra Leone, Slovakia, South Korea, Syria, Trinidad, Uganda, Ukraine, United Arab Emirates, United States, Venezuela, Vietnam, and Zimbabwe.

Our 2009 sample is by no means representative but provides interesting country coverage across several key variables.

While our ambition remains to cover all countries and regions of the world over time, the final set of countries for 2009 was chosen according to the following criteria:

Balance: Global Integrity seeks to ensure a basic geographic balance in its annual rounds of national-level country assessments. For 2009, the breakdown was as follows:

Sub-Saharan Africa: 8 (*Ghana, Kenya, Liberia, Malawi, Rwanda, Sierra Leone, Uganda, Zimbabwe*)

East and Southeast Asia: 4 (*China, Indonesia, South Korea, and Vietnam*)

Pacific: 0

Europe: 7 (*Bosnia and Herzegovina, Kosovo, Norway, Serbia, Slovakia, Macedonia, and Ukraine*)

Latin America: 5 (*Brazil, Colombia, Mexico, Nicaragua, and Venezuela*)

Middle East and North Africa: 6 (*Algeria, Jordan, Lebanon, Qatar, Syria, and United Arab Emirates*)

South and Central Asia: 6 (*Armenia, Azerbaijan, Georgia, India, Mongolia, and Nepal*)

North America and Caribbean: 2 (*Trinidad, United States*)

Budget: The total number of countries covered in 2009, 38, was the maximum number of countries we could afford to assess thanks to the generous support of our current donors.

Availability of experts: Perhaps the most important criterion that affects whether a country is selected is whether Global Integrity is able to recruit a sufficiently qualified research team in the country. We cannot, and do not, carry out country assessments where we lack confidence in our team. See additional details below on how country teams are recruited and selected.

The 2009 national assessment also represent our first attempt to adhere to a more standardized pattern of national coverage, which will entail biannual samples of a core group of 70 countries moving forward. For details, please visit: http://commons.globalintegrity.org/2009/06/global-integrity-report-2009-country.html.

Country Teams

Global Integrity combines the skills of leading journalists with expert social scientists and researchers to produce its comprehensive country assessments. In each country, Global Integrity retains a team of at least five to ten experts, working independently, to compile the material for the country assessments. A Global Integrity country team typically consists of:

- A lead reporter who reports and writes the Reporter's Notebook.
- A lead researcher who compiles the initial scores for the Integrity Indicators.
- At least three to five country "peer reviewers" (a mix of other in-country experts and out-of-country experts) who blindly review the raw data and reporting to provide unvarnished corrections, comments, and criticisms. Peer review comments on all data and reporting are published transparently alongside the original data and reporting, offering readers an alternative perspective.

The teams are coordinated from Washington via the Internet and phone. Until the public release of the country assessments, the researchers, journalists, and peer reviewers are unaware of the identities of other members of the country team. This is done to maintain the independence of the individual responses and avoid a peer-influenced consensus. All work is carried out by field teams via Global Integrity's custom-built Web-based data entry platform, MAGIC ("Multi-user Access to Global Integrity Content").

To identify appropriate researchers, journalists, and peer reviewers in each country, Global Integrity actively recruits potential team members through informal partnerships with several well-placed international NGOs that work in the anti-corruption and good governance field; those partners assist in publicizing our annual "Call for Experts." We also participate in international conferences and seek referrals from colleagues with whom we already work in various countries. Global Integrity then independently verifies the expertise and independence of prospective team members. For 2009, we received around 559 CVs from individuals who expressed interest in serving on a country team in one of our three distinct capacities (journalist, lead researcher, or country peer reviewer).

After a competitive internal selection process, the most qualified and available researchers, journalists, and peer reviewers in each country are identified, invited, and sent a contract with specific instructions on the scope of the work. The total number of country team members for 2009 was roughly 150 individuals. Interested readers can review the instructions sent to each team member through the following web link:

http://commons.wikispaces.com/MAGIC+Help+Desk

The Country Reports

The country assessments that make up the *Global Integrity Report* contain the following elements:

- **Corruption Timelines:** Compiled by Global Integrity researchers in Washington, these are unique political/historical timelines of significant corruption-related events at the national level. Designed as a quick reference resource, each timeline summarizes the main corruption-related events in the country from the early 1990s through present-day. Based largely on reliable English-language international and national media sources, the Corruption Timeline pulls together in an easy-to-read, chronological fashion what exists in the public domain about corruption in a particular country.

- **The Reporter's Notebooks:** Prepared by the lead in-country journalist, these are 1,250-word original essays analyzing the culture of corruption and state of governance in a particular country. This hard-hitting, impressionistic essay provides a snapshot of corruption in day-to-day life as well as the recent history and context regarding the most high-profile corruption cases dominating that country's media. Reporter's Notebooks are reviewed by the country's respective peer reviewers.

- **The Integrity Scorecard:** Scorecards are prepared and scored by the lead in-country researcher. Each country's 300-plus Integrity Indicators are aggregated to generate a country scorecard and the cross-country Global Integrity Index (see

additional details below). Integrity Scorecards are reviewed by the country's respective peer reviewers.

Note: "Country Facts" were eliminated as an element of each country assessment beginning in 2009.

Details on the Integrity Scorecard

The **Integrity Scorecard** for each country examines three concepts:

1. The **existence** of public integrity mechanisms, including laws and institutions, which promote public accountability and limit corruption.
2. The **effectiveness** of those mechanisms.
3. The **access** that citizens have to those mechanisms.

More specifically, indicators of **existence** assess the laws, regulations, and agencies/entities or equivalently functioning mechanisms that are in place in a particular country. Indicators of **effectiveness** assess such aspects of public integrity as protection from political interference; appointments that support the independence of an agency; professional, full-time staff and funding; independently initiated investigations; and imposition of penalties. Indicators of **citizen access** assess the ready availability of public reports to citizens, or publicly available information, within a reasonable time period and at a reasonable cost.

The Integrity Indicators are a unique instrument designed to provide a quantitative assessment of anti-corruption safeguards in a particular country at the national level. Carefully selected from a comprehensive review of the anti-corruption literature and other democratic governance sources, including Transparency International's National Integrity Systems framework, the Integrity Indicators are used to "score" the institutional framework that exists at the national level to promote public integrity and accountability and prevent abuses of power. For 2009, the Integrity Indicators were organized into six main categories and 23 sub-categories. They were:

1 Civil Society, Public Information and Media
1.1 Civil Society Organizations
1.2 Media
1.3 Public Access to Information

2 Elections
2.1 Voting & Citizen Participation
2.2 Election Integrity
2.3 Political Financing

3 Government Accountability
3.1 Executive Accountability

3.2 Legislative Accountability
3.3 Judicial Accountability
3.4 Budget Processes

4 Administration and Civil Service
4.1 Civil Service Regulations
4.2 Whistle-blowing Measures
4.3 Procurement
4.4 Privatization

5 Oversight and Regulation
5.1 National Ombudsman
5.2 Supreme Audit Institution
5.3 Taxes and Customs
5.4 State-Owned Enterprises
5.5 Business Licensing and Regulation

6 Anti-Corruption and Rule of Law
6.1 Anti-Corruption Law
6.2 Anti-Corruption Agency
6.3 Rule of Law
6.4 Law Enforcement

Generating an Integrity Scorecard

Each Integrity Indicator is scored directly by the lead researcher and substantiated as far as possible with relevant references and additional comments. The data is relayed from the field to HQ via the internet using MAGIC. There are two types of indicators: "in law" and "in practice." All indicators, regardless of type, are scored on the same ordinal scale of 0 to 100 with zero being the worst possible score and 100 perfect.

"In law" indicators provide an objective assessment of whether certain legal codes, fundamental rights, government institutions, and regulations exist. These "de jure" indicators are scored with a simple "yes" or "no" with "yes" receiving a 100 score and "no" receiving a zero.

"In practice" indicators address "de facto" issues such as implementation, effectiveness enforcement, and citizen access. As these usually require a more nuanced assessment, these "in practice" indicators are scored along an ordinal scale of zero to 100 with possible scores at 0, 25, 50, 75 and 100.

Lead researchers are required to provide a reference to substantiate each of their scores. This may be an interview conducted with a knowledgeable individual, a website link to a relevant report, or the name of a specific law or institution, depending on the particular indicator. Lead researchers are also offered the opportunity to include additional

comments to support their score and reference for a particular indicator. These are particularly useful in capturing the nuances of a particular situation, namely the "Yes, but…" phenomenon which is often the reality in undertaking this type of research.

Personality, language, and culture can all affect the interpretation of a particular indicator and the score assigned to it. To minimize this effect and maximize inter-coder reliability, Global Integrity provides researchers and peer reviewers with scoring criteria for every single Integrity Indicator. The scoring criteria anchor each indicator and sub-indicator to a predefined set of criteria. In essence, the scoring criteria guide the lead researcher by suggesting, "If you see X on the ground, score this indicator in the following way." For binary yes/no "in law" indicators, scoring criteria are provided for both "yes (100)" and "no (0)" responses. For "in practice" indicators, scoring criteria are defined for each of the 100, 50, and 0 scores with 25 and 75 deliberately left undefined to serve as in between scoring options. Scoring criteria for each indicator can be accessed via any of our online Integrity Scorecards by "hovering" one's mouse over a given indicator's scoring scale.

In summary, a given indicator or sub-indicator has the following elements:

- Indicator question, provided by Global Integrity
- Indicator scoring criteria, provided by Global Integrity
- Indicator score (either yes (100)/no (0) or ordinal scale of 0 - 100 with steps at 25, 50, and 75), assigned by the lead researcher based on:
  - References, provided by the lead researcher
  - Comments (optional), provided by the lead researcher
  - Peer review comments (optional), as provided through a double blind peer review process (see more on the peer review process below).

Example of Sub-Indicator Scoring Criteria

Public Access to Information > Sub-Indicator 13a
13: Is the right of access to information effective?
13a: In practice, citizens receive responses to access to information requests within a reasonable time period.

    ◻ 100   ◻ 75   ◻ 50   ◻ 25   ◻ 0

> **Scale 100 Criteria:**
>
> Records are available on-line, or records can be obtained within two weeks. Records are uniformly available; there are no delays for politically sensitive information. Legitimate exceptions are allowed for sensitive national security-related information.

> **Scale 50 Criteria:**
>
> Records take around one to two months to obtain. Some additional delays may be experienced. Politically-sensitive information may be withheld without sufficient justification.
>
> **Scale 0 Criteria:**
>
> Records take more than four months to acquire. In some cases, most records may be available sooner, but there may be persistent delays in obtaining politically sensitive records. National security exemptions may be abused to avoid disclosure of government information.

Sources: (required)

Identify two or more of the following sources to support your score.

☐ Media reports (identify the publication, author, date published, title, and website if available)

☐ Academic, policy or professional studies (identify the publication, author, date published, title, and website if available)

☐ Government studies (identify the publication, author, date published, title, and website if available) International organization studies (identify the publication, author, date published, title, and website if available)

☐ Interviews with government officials (identify the person by name, title, organization, date of interview, and place of interview)

☐ Interviews with academics (identify the person by name, title, organization, date of interview, and place of interview)

☐ Interviews with civil society or NGO representatives (identify the person by name, title, organization, date of interview, and place of interview)

☐ Interviews with journalists or media representatives (identify the person by name, title, organization, date of interview, and place of interview)

Description of sources (required).
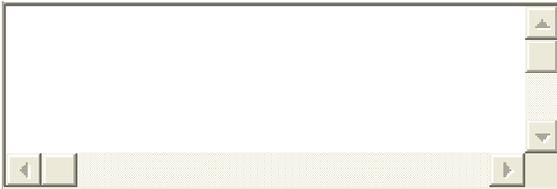
```
[text area]
```

Comments: (optional)

```
[text area]
```

Generating the Global Integrity Index: A Tool for Cross-Country Analysis

While it is important to stress that the strength of the Integrity Indicators is their utility as an in-depth diagnostic tool, Global Integrity is also able to generate scores across categories and countries and classify countries into an overall Global Integrity Index according to various performance bands. This may be a useful tool for those wishing to compare countries' overall performances against each another.

The Global Integrity Index groups countries into five performance "tiers" according to a country's overall aggregated score:

- Very strong (90+)
- Strong (80+)
- Moderate (70+)
- Weak (60+)
- Very Weak (< 60)

For the purpose of producing the Global Integrity Index, a simple aggregation method is used that parallels the process for generating indicator, sub-category, and category scores for each country. Original indicator and sub-indicator values are assigned by the lead researcher for the country (some of those scores may be adjusted following the peer review process; see more below). Each indicator score is then averaged within its parent subcategory, which produces a subcategory score. The subcategory score is in turn averaged with the other subcategory scores into a parent category score. Category scores are averaged to produce a country score which then falls into one of the above five groupings.
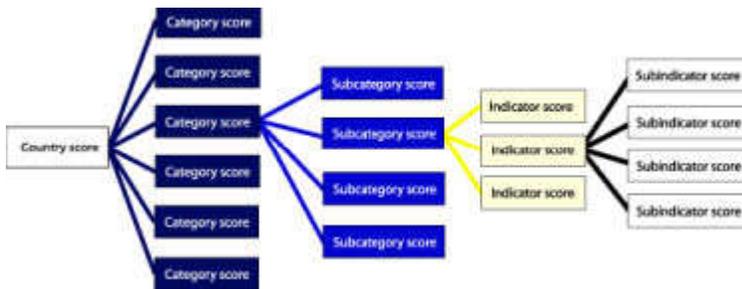
Because some aspects of governance and anti-corruption mechanisms are harder to measure definitively, some categories require a more complex matrix of questions than others. Thus, the categories are equally valued, even if some categories are derived from a more lengthy series of questions than others. Similarly, the subcategories are equally valued within their parent category.

In other words, each score (sub-indicators, indicator, and so on) is equally weighted with its peers addressing the same topic. However, indicators from different categories are not necessarily equally weighted. Our approach of using equally valued concepts and adding subordinate elements as needed has produced score weightings that reflect the six main conceptual categories evenly. Although we recognize the rationale for a non-equal weighting system (to give emphasis to issues of greater import), we have yet to develop a compelling defense for valuing certain categories, sub-categories, or indicators more importantly than others.

Peer Review Process

The importance of the peer review process cannot be overstated. Since Global Integrity utilizes an "expert assessment" approach to compile the Integrity Indicators and Reporter's Notebook, it is crucial that we employ quality control mechanisms to ensure that our data and reporting are as accurate and as balanced as possible. Individually contracted and carefully vetted "peer reviewers," selected for their independence and expertise in particular countries, are asked to blindly review both the raw Reporter's Notebook and the raw Integrity Indicators through MAGIC. The double-blind peer review process ensures that peer reviewers are unrestrained in their commentary, which likely adds to frankness when commenting on the draft country report. Peer review comments are used to interpret — and in some cases adjust — scores and reporting that peer reviewers identify as containing errors, bias, or out-of-date information. Indicator score adjustments follow certain rules and generally require repetition (i.e., similar comments from multiple peer reviewers) or solid referencing of a factual dispute. See further details on this below (Final Scores).

In reviewing the Reporter's Notebook for the country, peer reviewers are asked to consider the following:

- Is the Reporter's Notebook factually accurate?
- Is the Reporter's Notebook fair? Is anything misrepresented or unclear?
- Are there any significant events or developments that were not addressed?

Peer reviewer comments on the Reporter's Notebook are captured in narrative, paragraph form and are published anonymously alongside the final Reporter's Notebook.

For the Integrity Indicators, peer reviewers are asked to consider the following:

- Is the particular Indicator or Sub-indicator scored by the lead researcher factually accurate?
- Are there any significant events or developments that were not addressed?
- Does the Indicator or Sub-indicator offer a fair and balanced view of the anti-corruption environment?
- Is the scoring consistent within the entire set or sub-set of Integrity Indicators?
- Is the scoring controversial or widely accepted? Is controversial scoring sufficiently sourced?
- Are the sources used reliable and reputable?

The peer review process for the data scorecard, like that of the Reporter's Notebook, does not assign direct attribution to peer review comments. This ensures that peer reviewers are unrestrained in their commentary. Peer review comments on the country's data scorecard are published alongside the final scorecard and play an important role in final scoring adjustments prior to publication.

With regard to peer review feedback on the Integrity Scorecard data, peer reviewers are offered one of four standardized choices in responding to a given indicator or sub-indicator, using the above guidance to evaluate each data point:

1. *"Yes, I agree with the score and have no comments to add."*
2. *"Yes, I agree with the score but wish to add a comment, clarification, or suggest another reference."* Peer reviewers then provide their comment or additional reference in a separate text box which is published alongside the original data.
3. *"No, I do not agree with the score."* In this third case, peer reviewers are asked to explain and defend their criticism of the score and suggest an appropriate alternative score or reference.
4. *I am not qualified to respond to this indicator.*

For 2009, Global Integrity retained the services of approximately 100 peer reviewers for the annual around of country assessments, with some peer reviewers reviewing multiple countries.

Final Scores

Global Integrity takes full and final responsibility for the scores contained in the Integrity Scorecard for each country and the Global Integrity Index. These scores are generated following an elaborate and collaborative review process that includes balancing information from several (sometimes conflicting) sources while being guided by the master scoring criteria.

Following the peer review process, Global Integrity staff identify specific data points where peer reviewers have flagged problematic scores. The staff then engages the entire country team in a discussion of the issue in question and ultimately decides on appropriate changes, when necessary, to the original data based on the country team's feedback.

While Global Integrity makes every attempt to produce credible information, we welcome all feedback on the veracity and accuracy of our data. Please email Global Integrity with specific comments on indicator scores that you may not agree with, particularly with regard to factual accuracy.

Confidence Intervals

Beginning in 2007, Global Integrity began publishing margins of error for the top-level scores generated for countries as part of compiling the Global Integrity Index. We credit Dani Kaufman, Aart Kraay, and Massimo Mastruzzi of the World Bank and Brookings Institution for raising awareness within the research and practitioner community of the need to generate and acknowledge errors in any calculations of governance or corruption data (Kaufman, Kraay, Mastruzzi 2009).

The challenge in generating margins of error for the Global Integrity Index's country-level scores was that unlike a survey with a large sample of respondents, the Global Integrity approach relies on an expert assessment approach backed up by a transparent peer review process. Thus, we did not have a large "N" (sample size) that would allow us to use the ordinary techniques of variance and standard deviation to generate margins of error by exploring the variance of the data around the mean or median.

Instead, we relied on a count of the number of indicators for a country that were ultimately changed during the peer review process. We took those changes as an indication of the possibility that greater levels of error existed in the data for the country. In other words, the greater number of indicators changed because of the double blind peer review feedback, the greater the likelihood that other errors existed in the country's dataset, we reasoned.

However, we were careful to bear in mind that some countries would have a greater number of indicators flagged (and ultimately changed) because those scorecards were subjected to a larger than average number of peer reviewers. The number of peer

reviewers varies slightly from country to country each year. We did not want to automatically imply greater error for those countries that had more peer reviewers critically examining the data, a situation that, in our experience, almost always leads to a slightly increased number of indicators flagged for review.

We ultimately arrived at the following approach for generating margins of error for country-level scores:

Where:

N_i = number of indicators scored for the country (same for all countries)
M_{PR} = minimum number of peer reviewers for all countries
A_{PR} = actual number of peer reviewers for a country
C_i = number of indicators for the country changed as a result of the peer review process

Margin of error = $(C_i / N_i) * (M_{PR} / A_{PR}) * 100$

We then split the resultant margin around the actual country score to generate a +/- spread.

Example:

$N = 320$
$M_{PR} = 2$
$A_{PR} = 4$
$C_i = 23$

Margin of error = $(23/320) * (2/4) * 100 = 3.6$

If the country score were 74, we would publish the country score as 74 +/- 1.8 (1.8 equaling 3.6 divided by 2).

The (M_{PR} / A_{PR}) ratio serves as a "peer review factor" whose impact on the final margin of error diminishes as the number of actual peer reviewers for the country increases.

Implementation Gap

Beginning in 2007, Global Integrity began to publish what we have dubbed "implementation gaps" for each country. The implementation gap refers to the difference between the country's legal framework for good governance and anti-corruption and the actual implementation and enforcement of that same legal framework. We generate an implementation gap only at the country score level as opposed to the category and/or sub-category levels.

The implementation gap is created by first generating a "legal framework" score and an "actual implementation" score for each country. These two values are generated through two separate calculations. In each case, we perform the same aggregation technique described above for the country's Integrity Indicators *except* that we first remove either all "in law" or "in practice" indicators from the data set (for example, to generate the "legal framework" score, we first remove all "in practice" indicators from the country's data set and then perform the typical indicator aggregation technique described above). Once the "legal framework" and "actual implementation" scores have been generated, we simply subtract the "actual" score from the "legal" score to generate the implementation gap for the country.

Example:

56 (legal framework) – 26 (actual implementation) = 30 (implementation gap)

References:

Arndt, Christiane and Charles Oman. 2006. *Uses and Abuses of Governance Indicators* (Paris: OECD Development Centre).

Kaufmann, Daniel, Kraay, Aart and Mastruzzi, Massimo, Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2008 (June 29, 2009). World Bank Policy Research Working Paper No. 4978. Available at SSRN: http://ssrn.com/abstract=1424591

Medard, J.F. 2001. "L'evaluation de la corruption: approches et problemes," in *L'evaluation des politiques de developpement. Approches pluridisciplinaires* (Paris: L'Harmattan): 53-90.

Sik, Endre. 2002. "The Bad, the Worse and the Worst: Guesstimating the Level of Corruption," in *Political Corruption in Transition: A Skeptic's Handbook* (Budapest: Central European University Press): 91-113.