

# **The Politics of Governance Ratings<sup>1</sup>**

**Draft, 19<sup>th</sup> of February 2008**

*Do not quote without the authors' permission.*

**Christiane Arndt**

**PhD Candidate, Maastricht Graduate School of Governance**

**P.O. Box 616, 6200 MD Maastricht, the Netherlands**

**[christiane.arndt@governance.unimaas.nl](mailto:christiane.arndt@governance.unimaas.nl)**

**Charles Oman**

**OECD Development Centre**

**[charles.oman@oecd.org](mailto:charles.oman@oecd.org)**

Rapidly rising attention to the quality of governance in developing countries is driving fast growth in the use of governance “indicators”, for the purpose of both aid-allocation and investment decisions, and for academic analysis. Some International Organisations, and most prominently the World Bank, play a leading role, both in the supply and use of cross-country governance ratings. This paper attempts to explain i) the reasons for this role, ii) the problems associated with the most popular indicators produced by the World Bank, and iii) the reasons for the popularity and widespread misuse of these indicators. It argues that, while there will never be one perfect governance indicator, the production and use of more transparent governance indicators will better serve the needs of users and developing countries alike.

---

<sup>1</sup>This paper builds on the book “Uses and Abuses of Governance Indicators” by Arndt and Oman (2006). It bases many of its findings on interviews with donors, risk analysts, academics and OECD and World Bank staff who requested anonymity but whose assistance was invaluable. This paper benefited from enormously valuable suggestions and comments from Denis de Crombrughe and Chris de Neubourg at the Maastricht Graduate School of Governance, Matthew Andrews, Ricardo Hausmann, Pippa Norris and Lant Pritchett at Harvard University, Stephen Knack and Nick Manning and other, anonymous, commentators at the World Bank and Simon Kaja at the University of British Columbia. The authors would also like to thank participants of the workshop “Ranking and rating public services” in Oxford and participants of the seminar “The Politics of Governance Ratings” in Syracuse for their extremely useful comments and suggestions for improvement. The authors are solely responsible for the views expressed in this paper.

The last 15 years have seen a veritable explosion of interest in the quality of “governance“ in the developing world. This very growth has driven an equally significant increase in the use of quantitative governance indicators. The first section of this paper explains the reasons for the rising interest in governance. The second section introduces the most-widely used governance indicators and briefly explains their usage. The third section attempts to explain why international organisations (IOs) devise and use cross-country comparable governance indicators. The fourth section analyses the problems inherent in one of the most widely-used governance indicator datasets, the World Governance Indicators (WGIs) from the World Bank Institute. The fifth section attempts to explain the popularity of indicators like the WGI, and why they are frequently misused. Finally, the sixth section presents alternatives and ways to go forward. Many of the findings are based on interviews with development experts, undertaken between 2005 and 2007 (10 interviews with international banks, 6 interviews with bilateral donors, 12 interviews with staff in IOs (World Bank, UN, OECD), and 8 interviews with academics). Information was obtained from either or both questionnaires and (unstructured) interviews.

### **1. Why all the interest in governance?**

In recent years, four sets of phenomena have combined to drive the explosive growth of interest in the quality of governance - and with it the use of governance indicators.

A first set of phenomena is interlinked with the spectacular growth of international investment in developing countries over the last 15 years. International investors' major newfound interest in the quality of governance in developing countries is in part simply a reflection of the significant increase in the value of their assets exposed to risk in those countries.

A second set of phenomena responsible for the increased interest in the quality of governance in developing countries derives from the end of the Cold War. Throughout

the post-war period, the attitudes and behaviour of OECD governments and their national and multilateral aid agencies towards developing-country governments were dictated by the latter's position in the bi-polar world created by the Cold War. Only after the demise of the Soviet Union have these attitudes and behaviour become susceptible to real change – and increasingly so – with the perceived disappearance of the communist threat. A watershed was World Bank President James Wolfensohn's decision in 1996 to radically reverse the Bank's longstanding policy that it could not explicitly recognize or seek to address the acute problems of corruption in many of its borrowing countries, because local politics were outside the Bank's official mandate, and to give those problems a high priority. While World Bank lending to promote economic reforms fell by 14 per cent annually between 2000 and 2004, its lending to improve governance rose by 11 per cent annually during that period, so that by 2004 25 per cent of its lending was committed to law and public administration in borrowing countries (World Bank, 2004),

Growing perceptions in recent years of a relative failure or inadequacy of policy reforms widely undertaken in the 1980s and 1990s are a third set of phenomena driving the growth of interest in governance. The widespread policy shift at that time in favour of greater “market friendliness” in developing countries – that was actively encouraged by the multilateral financial organizations (sometimes referred to as “Washington Consensus”) that found themselves in a strong position to do so through lending conditionality – ultimately proved relatively disappointing (Easterly, 2002). There is today a growing understanding – including within the multilateral organizations and among staunch defenders of market-friendly policy regimes – that strong markets require good governance, and that poor local governance may go far to explain the relative reform failures of the 1980s and 1990s.<sup>1</sup>

Finally, a fourth set of phenomena driving the explosion of interest in governance has been the work of Douglass North and the New Institutional Economics of which he is a leading figure. This work has convincingly demonstrated the importance of a country's

system of governance for the country's success in terms of its long-term economic growth, enhancement of human welfare, and societal development (North, 1990, 2005).

## **2. The most-widely used governance indicators**

As international investors, aid donors and development analysts have increasingly come to understand the importance of governance, they have sought to render the concept operational for decision-making purposes. Following the maxim that 'you can only manage what you can measure', they have thus turned widely to using quantitative indicators to compare the quality of governance both among countries and over time.

Investors increasingly want back-up to traditional, macroeconomic-based, country-risk indicators that have failed to predict costly financial crises. They base their investment-location decisions on their perceptions of governance and these perceptions are increasingly shaped by governance indicators. Bi- and multilateral donors want to identify and reward developing countries that are improving their governance – and punish those that aren't. Many of these agencies also claim to use governance indicators to increase the transparency and consistency of their criteria for determining recipients' aid eligibility. The US Millennium Challenge Corporation and the World Bank's International Development Association each year allocate billions of dollars of aid on the basis of governance indicators. Finally, development analysts need governance indicators for their quantitative governance studies and journalists use them to report 'horse races' between countries. Thus, most users of governance indicators use them to compare the quality of governance across countries and over time.

Developing countries' own governments, business associations, NGOs and other such "internal stakeholders" who wish to bring about improvement in local governance rarely use governance indicators. They would need indicators that identify specific governance problems and monitor progress in reform with much greater precision than most currently

available governance indicators. Recently, much progress has been made in the construction of governance indicators to better satisfy the needs of developing countries.

Though there are hundreds of governance indicators, a few dominate the market. Investors, donors, journalists and development analysts rely very heavily on a small selection of composite, mostly perception-based indicators. Perception-based indicators rely on the perceptions of households, business people, civil servants or experts of the quality of governance. Facts-based indicators are derived from objective facts such as data on the existence of regulations that make it more difficult to fire employees. While the data used to construct facts-based ratings are objective, the ratings themselves are often highly subjective in both the choice of facts and the interpretation of what these facts means (e.g. “Is a low level of labour standards a hallmark of good or bad governance?”)

The most popular indicators compile a large amount of information which they reduce to a single number for a given country in a given year. It is precisely reducing the huge amount of information to a single score that makes them so attractive for decision-making purposes. Among the most widely-used are the so-called “World Governance Indicators” (WGIs) produced since 1996 by Daniel Kaufmann and his team at the World Bank Institute, Transparency International’s “Corruption Perceptions Index” (CPI) and the facts-based “Doing Business Indicators” from the World Bank and the International Finance Corporation (IFC) Also widely used (and incorporated into the “WGIs”) are the International Country Risk Guide (ICRG), Freedom House’s “Political Rights” and “Civil Liberties” indicators and the Governance Indicators from the World Bank’s “Country and Policy Institutional Assessment” (CPIA). (see Hood *et al*, 2007, and Arndt and Oman, 2006<sup>2</sup>)

The World Bank therefore produces three of the most popular and influential governance ratings. For its own aid-allocation, the World Bank relies on the CPIA indicators which

are based on World Bank staff assessments of the quality of governance as well as economic and social policies. The six WGIs are the result of aggregating the governance ratings of more than thirty expert assessments (which include indicators of the CPIA) , firm and household surveys. While the World Bank itself does not use the WGIs in an official way, other organisations do: The US Millennium Challenge Corporation (MCC), for instance, relies heavily on them for allocating billions of dollars to developing countries.

### **3. Why do International Organisations produce governance ratings?**

IOs produce and use cross-country comparable composite governance indicators to create incentives in developing countries to improve governance, to enhance the transparency and objectivity of their aid-allocation decisions, to undertake quantitative governance research, and for institutional reasons. Whether the most-widely used governance indicators fulfil these expectations is the subject of an ongoing debate. This section focuses on cross-country comparable ratings and not on country-tailored governance indicators. The United Nations Development Programme (UNDP), for instance, is involved in country-specific governance assessments, but explicitly refrains from undertaking assessments that compare countries' governance performance, with the exception of governance related to gender equality.

#### *Incentives for developing countries to improve their governance*

Governance ratings can create incentives in developing countries to improve their governance in two ways: Through i) “naming and shaming” and “horse races”, and through ii) aid-allocation conditionality. To illustrate, Transparency International's Corruption Perceptions Index (CPI) was successful in drawing attention to the problem of corruption and in putting it up-front on the agenda and the WGIs triggered a “horse race” in the media between developing countries. Governments seeking external funding also increasingly pay attention to governance ratings.

However, some developing countries' governments and scholars now increasingly tend to resist international governance ratings, according to the experience of some World Bank and UNDP staff members interviewed for this paper. They challenge the indicators for normative, legitimacy and methodological reasons. Some developing countries express their distrust to indicators that they resent as *pro*-liberalisation and *pro*-market and therefore as a continuation of the Washington Consensus. They accuse governance indicators to reflect a Western or Anglo-Saxon type of governance ideal that is not helpful to understand the governance reality in their countries (Girvan, 2002). Krishna Guha reports in the Financial Times (2007a) that nine of the World Bank's executive directors, including from China, Russia, Mexico and Argentina, complained about the most recent WGI ratings: "Some governments argue that the bank is in no position to preach on governance, following its own crisis over former president Paul Wolfowitz's handling of a secondment package for his girlfriend. Others, including China, question whether the World Bank should be involved at all in rating countries on metrics such as accountability" (Guha, 2007b).

Some World Bank staff reported that the ratings are harmful for their relations with local governments. To illustrate, one government laughed at the WGIs and the World Bank lost credibility in the country. Another illustration is the non-acceptance of low CPIA ratings by developing country governments. This explains partly why countries' ratings are too high, according to some experts interviewed: A World Bank country director has both an incentive to have good relations with the government as well as to obtain resources for the country. Both depend on a good CPIA rating on which the country director has a high influence.

The lack of "actionability", transparency and comparability over time of the most-widely used governance indicators is a further hindering factor in pressuring developing countries into improving their governance. Several World Bank economists, who work in country offices, said that they were unable to fulfil the requests of developing countries

governments to explain to them their ratings on the WGIs. Many of the IO staff, interviewed, who are working directly in developing countries, said they do not use broad governance indicators like the WGIs on the ground: Some staff felt that it is difficult to start a policy-oriented debate in their country on the basis of broad composite indicators and some resented the stigmatization of countries through international governance ratings. Others found the indicators useful to raise awareness of governance problems in countries where the government is not very willing to reform and to start a dialogue with the government. Both more specific cross-country comparable indicators as well as in-depth assessments in a country - for instance the World Bank Institute's Governance Diagnostic Surveys<sup>3</sup> - were cited by IO country office staff as useful for governance reform.

#### *Higher transparency of aid-allocation decisions*

Donors are increasingly aware of the lack of good governance in their own aid-allocation processes and, therefore, consider governance indicators a tool to enhance the objectivity and transparency of those decisions. Paradoxically, the indicators they use lack just that transparency and contain hidden biases, as will be explained in the next section of the paper.

#### *Quantitative governance analysis*

IOs also produce and use governance indicators for quantitative governance analysis to enlighten the governance-development relationship. Studies that ascertain evidence for a positive impact of good governance on economic growth or on the efficient use of development aid serve as a justification for spending money on governance reform in developing countries (for instance Burnside and Dollar, 1997, Knack and Keefer, 1995 or Kaufmann and Kraay, 2002). These studies were crucial in the change of agenda towards governance reforms in developing countries.

Yet researchers' use of governance indicators is widely plagued with pitfalls, as the nature and limitations of composite governance indicators unfortunately weaken the rigour and the credibility of many studies. The next part will illustrate some technical and conceptual problems with the example of the WGIs which are used frequently in academic studies.

### *Institutional reasons*

A number of institutional reasons reinforced the rise of governance indicators in IOs. First of all, governance indicators give institutions a profile. It is a product that is very visible within and outside the organisation. Secondly, if a leading actor adopts a new policy, other actors are likely to follow independent of its efficiency (DiMaggio and Powell, 1991). The World Bank produced and used governance indicators and many other development agencies and Ios like the the OECD and the Asian Development Bank followed. They followed, because i) governance indicators are “trendy” and ii) because they aimed at improving upon existing measurements or complementing them with different measurements. Thirdly, visible work on the increasingly important issue of measuring governance can help units within an international organisation to obtain a more important role. Dunleavy's (1991) bureau-shaping model predicts that bureaucrats would prefer to work in small, elite agencies close to political power centres and doing interesting work, rather than to run large-budget agencies with many staff but also many risks and problems. To illustrate, the World Bank Institute rose through its governance assessments and its governance expertise from a training institute to having more strategic importance and the Doing Business Indicators made the IFC more prominent. It is precisely the high level of usage and media coverage of their indicators that benefited the World Bank Institute and the IFC.

The increase in the number of producers of indicators has had a positive effect on the quality and diversity of governance indicators available. To illustrate, in response to a previous lack of data several new initiatives (e.g. DIAL's household surveys) produce

indicators that reflect the experiences of the poor, and other initiatives try to produce more transparent indicators (e.g. Global Integrity). However, it is difficult for new initiatives like Global Integrity and DIAL to gain attention, because the most-widely used indicators are well-established and dominate the market. The success of the most-widely used indicators has created interests which make it more difficult to move on to new developments. The development of more (context) specific indicators would shift more responsibility to the staff working “on the ground”, away from the headquarters.

The downside of the “measuring governance trend” is that organisations spend large amounts of money on assessing governance, while the funds might, in some cases, be better used elsewhere. Jak Jabes (2002), Director of the Governance and Regional Cooperation Division at the Asian Development Bank, cautioned: “Governance indicators are ultimately useful if they can be utilized to mobilize the leaders of a country to bring about change. If, as I unfortunately think is happening, indicators become an end in themselves, then they end up becoming part of dust-gathering reports and detract attention from a less complicated yet more practical and productive approach”.

#### **4. The World Bank Institute’s “World Governance Indicators”**

The six WGIs, namely “Voice and Accountability”, “Political Stability”, “Government Effectiveness”, “Regulatory Quality”, “Rule of Law” and “Control of Corruption”, are the most popular governance indicators. Publicly available and easily accessible on the web, they are widely-used to compare the quality of governance over time and across countries, for aid-allocation decisions, for risk ratings, for academic analysis and for media articles. They summarize data from more than 30 expert assessments and household and firm surveys, and are available from 1996 to today for more than 200 countries.

According to one of the authors<sup>4</sup>, the WGIs were created in response to four inter-related concerns: One was the apparent lack of robustness of cross-country comparisons using

different individual data sources, especially when the different sources led to different conclusions. Second was concern about how to interpret cross-country differences and their statistical and practical significance. Third was concern that it is difficult to compare results from regional surveys with broader cross-country surveys. A fourth concern was to find a way to produce useful overarching, integrative or summary indicators, given the large and growing diversity of individual sources cropping up in different pieces of research and in policy debates.

While the WGIs are a very useful first snapshot of experts' perceptions of a country's quality of governance, users often fail to take into account their limitations for cross-country and over time comparisons. This section will briefly explain five interrelated concerns with the indicators that users are often not aware of: i) lack of transparency, ii) absence of an underlying theory of good governance, iii) hidden bias, iv) lack of comparability over time and v) lack of actionability. The interested reader can find a more exhaustive discussion in Arndt and Oman (2006).

### *The Transparency Paradox*

Some donors have come to give high importance to the use of relatively untransparent indicators like the WGI or the CPI in, paradoxically, trying to give transparency to their aid-allocation decisions. The reasons for the WGI's lack of transparency are threefold: i) high complexity, ii) lack of reproducibility, iii) lack of underlying theory of governance.

First of all, the sheer number and diversity of indicators, produced by others, that are incorporated into a single WGI make it very difficult to understand a country's rating, as illustrated in Figure 1 for "Regulatory Quality" in Bangladesh 1998 and 2005. At the tip of the pyramid is the rating for Regulatory Quality of Bangladesh. Underneath are the sources - surveys or expert assessments - from which that rating is calculated. Most sources are themselves composite indicators, constructed from experts', business' and households' answers to survey questions or checklists (lower part of the pyramid).

Understanding a change in a country's rating over time or a cross-country difference is complicated, because the composition and weights of primary sources vary across countries and over time.<sup>5</sup> To illustrate, Knack (2006) explains that the WGI "Control of Corruption" uses 23 different combinations of sources for the Eastern European and Central Asian countries and that there are only four pair of countries whose ratings are based on a common set of sources. This difficulty is all the greater because some of the underlying indicators used as inputs are themselves very broad and imprecise or lacking transparency. To illustrate, it is not clear what exact criteria the Economist Intelligence Unit (EIU) uses for rating countries on the indicator "Excessive Protections", one of the sources for Regulatory Quality in Bangladesh (see Figure 1). When are protections necessary and when are they excessive according to the EIU?

Secondly, the WGIs are not replicable. Many of the indicators underlying each source's ratings, the 2nd layer in the pyramid, are not disclosed to the public (see Figure 1)<sup>6</sup>. Kaufmann *et al.* publish only the overall rating of each primary source in rescaled form.

Thirdly, the WGIs lack any underlying theory of good or bad governance that would explain the meaning of the indicators.

#### *Absence of an underlying theory of good governance*

It is important to remember that all rating systems are normative by their very nature. They imply a judgement, if they rate one country better than another country. One might, therefore, expect the WGIs to be based on a normative definition or concept of what constitutes good or bad governance which determines the countries' ratings. However, there is no such normative concept or theory underlying the WGI ratings. The authors define these ratings simply as "the traditions and institutions by which authority in a country is exercised". They further disaggregate the channels into different components, yet they do not explain how to distinguish between good or bad governance.

On the contrary, it's the compilation of all the ratings from different sources brought together in a single "WGI" which gives normative meaning to the concept of governance. But nobody, including the authors themselves, can define these norms. In addition, the norms differ from year to year and from country to country, because the sources for the WGIs, as for the CPI from Transparency International, differ across countries and over time.

Melissa Thomas (2007) finds no evidence for the construct validity of the WGIs. She concludes that the "indicators stand as an elaborate and unsupported hypothesis" and that "accordingly, reliance on the indicators is premature" (p.2)

Furthermore, rating countries on "true governance" implies that there is only one, "true", system of good governance. One may wonder about the nature of the opposite, "untrue governance," and whether a single governance system should be imposed on the developing world. Meisel (2004) explains why "Anglo-Saxon glasses" do not help to explain the rapid growth in the Asian economies and offers a new model of "Governance Cultures", helpful for developing countries today.

### *Hidden bias*

The problem is not so much that there is bias in a given WGI indicator. Every governance indicator that ranks countries will be biased, because every rating of a country is based on norms and preferences that are not necessarily shared worldwide. The problem is that the biases remain largely hidden from users, as the more than thirty sources, many of which lack transparency themselves, determine the normative meaning of the indicators.

Regulations to protect the environment and labour rights tend to be interpreted as 'unfriendly' to markets by the indicators that carry the most weight in the aggregation process. Whereas one may or may not agree with this interpretation, the problem is that most users are not aware of this bias. For example, Kaufmann *et al.* include the indicator

“Environmental regulations hurt competitiveness” from the World Economic Forum’s Executive Opinion Survey, but they do not include several questions that give high ratings to countries with a high standard of environmental protection. Users are generally not aware of this omission.

Another illustration is the low weight given to household surveys relative to the weights of expert assessments and firm surveys. To illustrate, Gallup’s World Poll that asks citizens about their exposure to crime gets zero weight for the WGI “Rule of Law”, whereas Global Insight Business Risk and Conditions, a commercial business information provider from Boston that measures the crime risk to businesses, gets the third highest weight<sup>7</sup>.

Two reinforcing factors lie behind the low weights: first of all, there are fewer cross-country comparable household surveys and they provide data for fewer countries, because they are expensive and because there are few opportunities to exploit the data commercially. Kaufmann *et al.* use most cross-country comparable governance indicators available and commercial interests drive a major part of this supply of governance indicators. Secondly, while the authors are not responsible for the low number of household surveys available, they could have corrected for this selection bias in the aggregation process. Instead, they chose to use an aggregation technique<sup>8</sup> which gives less weight to perceptions that diverge from the dominant majority than to perceptions that conform to the majority view. Household surveys diverge from the majority of expert perceptions, and therefore, get low weights. The reason for this divergence is not necessarily that either household surveys or expert assessments are wrong, but could also be that they assess different aspects of governance. To come back to our previous example, the crime rate and protection (or abuse) by the police and judiciary are not necessarily at the same level for the poor parts of a city and the diplomatic and business parts. However, Kaufmann *et al.* assume that the discrepancy between Gallup’s rating

and the majority of expert ratings is due to its high measurement error and discard the possibility that it measures something different.

#### *Lack of comparability over time*

The WGI should not be used for direct comparisons of levels of governance over time. An actual change in the level of governance is only one of four possible reasons for a change in a country's rating on the WGIs. A second possible reason is a change in the sources' perceptions of the quality of governance without an actual change in the quality of governance. For instance, in December 1996, a string of corruption cases hit Charles Haughey, who had been prime minister in Ireland from 1979 to 1992. The International Country Risk Guide, one of the sources for the WGIs, downgraded Ireland several times in the months that followed and during which more and more evidence of the scandals appeared. It is unlikely that the level of actual corruption had changed in Ireland, but the ICRG updated its rating based on newly available information. A third possible reason for a change is that other countries' governance ratings changed. Since Kaufmann *et al.* scale the indicators in such a way that the global average and standard deviation are always the same, changes in one country's rating automatically change the ratings of other countries, other things equal. A fourth possible reason is a change in the number and composition of sources from one year to the next. Changes in the components of a composite indicator over time are not necessarily problematic for its comparability over time. The components of the basket of goods for the Consumer Price Index, for instance, change frequently. However, this change is motivated by a substantive reason: the index purports to measure the increase in costs of what people actually consume and that changes over time. To illustrate, to own a computer was exceptional for a household ten years ago and it is standard nowadays. For the WGIs, changes in sources over time are not driven by substantive reasons, but by the availability of more data.<sup>9</sup>

#### *Lack of actionability*

The WGIs give little guidance to local stakeholders or others as to what concrete actions they can or ought to take to improve the quality of local governance. For instance, an indicator for Rule of Law tells us “how secure business people feel about their property,” but it tells us very little about what makes them feel that way. The indicators are used mainly as a decision tool for outsiders, who wish to judge or compare countries, but are not action-oriented for stakeholders in the countries that are being judged and compared.

### **5. Why are composite cross-country comparable governance indicators so popular despite their shortcomings? Why are they misused?**

The last paragraph highlighted, with the example of the WGIs, the limitations of the most popular governance indicators, all of them highly aggregated summary measurements, and Arndt and Oman (2006) showed that these indicators are often not correctly used. Why are these indicators so popular despite their shortcomings? Why do donors, investors and analysts use such indicators for purposes they are not designed for?

The following four major reasons lie behind the market domination of a small number of composite indicators and their misuse: i) users do not understand the limitations of the indicators, ii) complex composite indicators meet an important demand for summary measurements of the quality of governance, iii) users follow other users, and iv) users perceive that there is no alternative.

#### *Incomplete understanding of indicators' limitations*

Users do not fully grasp the limitations of the indicators they use and have therefore wrong perceptions of the reliability and content of indicators. In particular, they are often not aware of problems with the comparability over time and across countries, and of the normative content of indicators. (To illustrate, a recent paper from a renowned university refers to the WGIs and Freedom House indicators as “objective”<sup>10</sup>). The reasons for this misunderstanding are multiple and intertwined:

Firstly, the complexity of composite indicators makes it very difficult to understand their limitations. Many users do not have the time or do not see the need to see through this complexity, in order to deeply understand the way the indicators they use are constructed. Furthermore, understanding the complicated methodological background paper of, for instance, the WGIs requires not only time, but also an advanced level of statistical expertise that many users, often generalists or experts in other fields, lack. Some users might trust in the reliability of the indicators precisely because they are impressed by the equations and tests in the background papers that they do not fully understand.

Secondly, some users think that relying on a large number of different sources eliminates the bias inherent in individual sources. The normative content and bias is often hidden and users are therefore not aware that they may not share the norms inherent in the indicators.

Thirdly, non-statistical experts often tend to believe that numbers are facts. As soon as a governance assessment is quantitative, the assessment is believed to be objective. Numbers reassure and create the impression of indisputable facts.

Fourthly, misleading advertising contributes to a considerable misperception and therefore misuse of governance indicators. While Transparency International carefully advertises its CPI as “a snapshot of the views of business people and country analysts, with less of a focus on year-to-year trends”, the World Bank Institute advertises its WGIs as “reliable measurements of governance”. Presentations and its website also create the misleading impression that the views of ordinary citizens are well-represented, making the indicators particularly attractive to donor agencies that seek to improve the living conditions of the poor. To illustrate, Kaufmann *et al.*(2007) heavily emphasize the recent inclusion of the source “Gallup World Poll “, a cross-country household survey available for a large number of countries. What they do not emphasize in the paper is that Gallup’s

World Poll gets zero weight on two WGI, marginal weight on two other WGI and provides no data for the remaining two.

Fifthly, users often do not fully understand or ignore warnings of the producers to make cross-country and over-time comparisons only with due caution. To illustrate, Kaufmann *et al.* very carefully flag many cross-country differences or over-time changes in the WGI as “statistically insignificant”, a warning that is often ignored, in particular in newspaper articles. Journalists may ignore these warnings on purpose, because taking them into account would make their article too complicated and less interesting, or they may not be familiar with the concept of statistical significance themselves.

#### *Supply meets demand for summary measurements*

There is a high-demand from development agencies, ministries, IOs, international banks and the media for summary measurements of the quality of governance with broad country coverage. Summary measurements are convenient for busy decision-makers. They are quick to look at and therefore help to make quick decisions, compared to looking at hundreds of different indicators<sup>11</sup> or qualitative reports.

Using a summary measurement of most available sources frees decision-makers and development researchers from i) getting lost in the complexity of governance and ii) making explicit normative choices that could render them vulnerable to criticism. To illustrate, Steven de Walle (2007) shows that concepts used in evaluating the performance of public services appear to be neutral, but are generally value-laden. Protecting public-service values such as universal access to services, including remote areas, can come at the sacrifice of efficiency and there is no neutral answer to what is more important. Similarly, most donors or academics agree that the quality of regulations should be good. The disagreement arises as soon as they try to specify what regulations are a sign of good quality: Are environmental regulations deemed ‘good’ because they protect the environment and the health of the population, or should they considered to be

‘bad’, because it puts businesses under too much financial and administrative strain to comply with them? Most people also agree that governments should protect the freedom of the press, but they may disagree on the specific weight that different aspects of press freedom should carry in a “Press Freedom Index”. To illustrate, what weight should be given to an indicator assessing the number of journalists who were killed in a given country, over an indicator that assesses ownership concentration in the media sector?

Faced with the complexity of governance, difficult normative choices and high time-pressure, it is not surprising that decision-makers and development analysts often opt for well-known and widely-used summary measurements.

One of the World Bank staff members interviewed for this paper said that such measurements help to “avoid the hard part of development, which is working out where and how to assist” and that providing such governance indicator rankings to governments “conveys no more information than just saying to them ‘you must develop’”.

### *Media*

The media extensively uses and publishes ratings, because ratings sell. US News & World Report for instance generated a lot of publicity over its annual ranking of colleges and universities and continues to publish the ratings, despite protests of American Universities about the arbitrariness of these indicators. The controversy is similar to the controversy over the CPI and the WGIs.

The more the media reports on countries’ positions on the CPI or the WGIs, the more their use is reinforced. The reason is threefold: Firstly, potential users learn about the CPI or the WGIs in the media and may not know about other initiatives. Secondly, even if they know about other initiatives, users may trust these indicators more than unknown indicators. Thirdly, the producers of these indicators can use these media reports both internally and externally to justify and promote their work.

New initiatives like Global Integrity face not only an uphill battle against other indicators that have been around longer, but also against indicators that are available for larger number of countries. As a World Bank economist phrased it, “a better or more meaningful index can be crowded out by a less meaningful one that has broader country coverage”.

Political reasons may also explain why some indicators are more promoted and passed on to the media. To illustrate, in the field of corruption, the WGI “Control of Corruption” which focuses on bribery and other illegal corrupt activities, is professionally promoted. It is interesting that the World Bank Institute’s work on legal corruption (Kaufmann, 2004) is barely promoted and therefore barely known. It takes into account legal ways (lobbying or party-financing) for special interest groups to influence laws and regulations in a country. These are more common in OECD countries than bribery to influence the government in favour of special interests. The Corporate Legal Corruption Index ranks the United States, a powerful and highly influential member of the World Bank, only 53th in 2004 (the lower the rank, the better), while it scores better than 93% of countries on the Control of Corruption Indicator 2004.

#### *Herd behaviour*

Governance Indicators obtain authority through the number and importance of users. Other users, therefore, often serve as a justification for someone’s own use. To illustrate, the authors of the IMF Country Report on Albania (2006) first describe the limitations of the WGI and then proceed to say “However, the government (...) recognises that the World Bank governance indicators are generally accepted internationally.”

#### *Perceived absence of an alternative*

Many users perceive that there is no alternative to the most-widely used governance indicators. The reason is threefold: first of all, many alternative governance indicators are

only available for a limited number of countries, whereas indicators such as the WGI draw their utility precisely from the fact that they use all available information to obtain a rating for almost all countries in the world. Secondly, users are likely to think that the more sources a composite indicator uses, the more reliable its rating will be. Thirdly, users may not even know about innovative approaches to quantify the quality of governance, because the media focuses its attention on established indicators like Transparency International's CPI.

## **6. Alternatives**

There are several alternatives to relying heavily on a few summary measurements of the quality of governance. In the short-run, users can find advice and guidance about the measurements that will most likely suit their purposes in governance indicator guides and inventories (IADB, 2007, UNDP, 2007, Metagora, 2007, World Bank Institute, 2007, Besancon, 2003). These are useful: i) to gain awareness of and profit from the wealth of existing data and ii) to understand the indicators' limitations and hence use them correctly. The Inter American Development Bank's (IADB) web tool, for example, provides information for 400 indicators about the methodology used to build them and the implications this has for the indicators' reliability, validity and suitability for making comparisons across countries and over time. In the short-run, producers can also improve the way they advertise their indicators and improve their transparency to better convey to users the meaning and limitations of their indicators.

In the medium-to long run, the trend towards more specific, meaningful and actionable indicators will continue. While we do not have at our disposal a theory of governance able to guide the meaningful construction of indicators of the overall quality of governance, some aspects of governance are well understood and indicators can therefore be based on theory. These more specific indicators are often also composite indicators, but differ from opaque summary measurements in that both the choice of the indicators and the weighting is transparent and based on theory. An illustration is the multi-agency

“Public Expenditure and Financial Accountability Program” (PEFA) that rates countries on their public management performance in a transparent manner, publishing the results along with detailed explanations for countries’ ratings.

Quantifying the quality of governance in a transparent and specific way and trying to make indicators “actionable” forces development analysts to face the limits of their knowledge. One major motivation for international organisations to develop “actionable” indicators is to induce governments to reform in order to improve their rating on these indicators. Developing indicators that suggest specific reforms requires a deep understanding of what governance reforms are conducive to development. According to some development experts interviewed the lack of consensus on what governance reforms are necessary in developing countries and to what extent these reforms are context-specific, might precisely explain the current lack of a large number of cross-country comparable actionable indicators (see also Grindle, 2007). Some experts expressed their fear that an overambitious good governance agenda, reflected in a broad set of good governance indicators, might render a developing country’s government immobile and divert its attention from more urgent problems. Several development experts argued for more context-specific governance assessments that support governments in setting priorities for governance reform and monitor progress over time (e.g. World Bank Institute Governance Diagnostic Surveys, UNDP governance assessments). Some experts also expressed the need for more qualitative work, to better understand developing countries’ governance systems (e.g. Meisel, 2004).

## **Conclusion**

Some International Organisations, and most prominently the World Bank, play a leading role both in the use and the supply of governance indicators. The most popular governance indicators are perception-based composite indicators that summarize a large amount of data and are available for a large number of countries. IOs devise and use composite cross-country comparable governance indicators to create incentives in

developing countries to improve their governance, to enhance the transparency and objectivity of their aid-allocation decisions, to undertake quantitative governance analysis and for institutional reasons. However, the limitations of the most-widely used indicators with respect to comparability over time, transparency and actionability and the increased resistance from developing countries towards “Western governance indicators” puts into question the fulfilment of these high expectations, according to many of the development experts interviewed.

While the most popular indicators are very useful in gaining a first idea of how experts perceive a country’s quality of governance, they are also widely used for aid-allocation decisions, country risk analysis and media articles that praise or criticize a government – with too little attention paid to the indicators’ limitations. The more users rely on the same indicators, the more they become “generally accepted internationally”, which spreads their use further. The high visibility of a few indicators in the media accelerates this process of concentration, to the detriment of new indicators that might be more meaningful. Another reason for the popularity of a few indicators and their widespread misuse is that decision-makers demand a summary measurement for cross-country and over-time comparisons and perceive that there is no alternative to the most popular indicators.

Many users are often not aware of the wealth of information available and that they can find advice and guidance about the measurements that best suit their purposes in governance indicator guides and inventories (IADB, 2007, UNDP, 2007, Metagora, 2007, World Bank Institute, 2007, Besancon, 2003). New, more specific and transparent initiatives have emerged in recent years. Some provide comparable data for a large number of countries and others are more context-specific. What they have in common is that they do not claim to assess the quality of overall governance, but focus on specific and well-defined aspects of governance and, therefore, give guidance to developing countries on how to improve their rating. Until we have at our disposal a theory of

governance able to guide the construction of meaningful indicators of the overall quality of governance, the development community should focus on such specific and well-defined indicators. As with all indicators, these new indicators have their limitations; however, since they are transparent, their limitations are also transparent, hence reducing the danger of misinterpretation and misuse.

## Reference list

- Arndt, C. and Oman. C. 2006. *Uses and Abuses of Governance Indicators*. Paris: Development Centre, OECD.
- Besançon, M. 2003. "Good Governance Rankings: The Art of Measurement.". World Peace Foundation Reports, vol. 36, Cambridge: WPF.
- Burnside, C. and D. Dollar. 1997. "Aid, Policies, and Growth." Washington, D.C: World Bank.
- Cornia, G. and V. Popov ed. 2001. *Transition and Institutions: The Experience of Gradual and Late Reformers*. Oxford: UNU/WIDER, Oxford University Press.
- DiMaggio, P. and Powell, W. 1991. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organisational Fields" in *The New Institutionalism in Organisational Analysis*. London and Chicago: The University of Chicago Press.
- Dunleavy, P. 1991. *Democracy, Bureaucracy and Public Choice: Economic Explanations in Political Science*. New York: Harvester Wheatsheaf.
- Easterly, W. 2002. *The Elusive Quest for Growth*. Cambridge, Massachusetts: MIT Press.
- Girvan, N. 2002. Problems with UNDP Governance Indicators. Available online at <http://www.southcentre.org/info/southbulletin/bulletin41/bulletin41-04.htm> (accessed June 2007).
- Global Integrity. Assessments and Data. Available online at <http://www.globalintegrity.org/data/index.cfm> (accessed June 2007).

- Grindle, M.S. 2007. "Good Enough Governance Revisited." *Development Policy Review* 25(5): 553-74.
- Guha, K. 2007a. "Countries Unhappy at World Bank Ratings." *Financial Times*, July 10, 2007.
- Guha, K. 2007b. "World Bank Directors Challenge Report." *Financial Times*, July 12, 2007.
- Hood, C., Beeston, C. and Dixon, R. 2007. "Rating the Rankings: Assessing International Rankings of Public Service." Paper prepared for the *IPMN Workshop, Ranking and Rating Public Services*. Worcester College, Oxford, 7-9 August 2007.
- IMF. 2006. "Albania: Poverty Reduction Strategy Paper—Annual Progress Report.", available online at <http://www.imf.org/external/pubs/ft/scr/2006/cr0623.pdf> (accessed June 2007)
- Inter-American Development Bank. 2007. Datagob: Governance Indicator Database. Available online at <http://www.iadb.org/datagob/>. (accessed June 2007).
- Jabes, J. 2002. "On the (F)utility of Governance Indicators: Lessons from Countries in Transition", Speech given as part of the "Assessing Good Governance Panel" at the *International Institute of Administrative Sciences Second Specialized International Conference*. New Delhi.
- Johnson, D. and Zajonc, T. 2006. "Can Foreign Aid Create an Incentive for Good Governance? Evidence from the Millennium Challenge Corporation." Harvard University, Working Paper.
- Kaufmann, D. 2004. *Corruption, Governance and Security: Challenges for the Rich*

- Countries and the World. Available online at <http://www.worldbank.org/wbi/governance/pubs/gcr2004.html>. (accessed October 2007).
- Kaufmann, D., A. Kraay and M. Mastruzzi. 2006. "Governance Matters V: Governance Indicators for 1996-2005." Washington, D.C.: World Bank Institute.
- Kaufmann, D., A. Kraay and M. Mastruzzi. 2007. "Governance Matters VI: Aggregate and Individual Governance Indicators for 1996-2006." Washington, D.C.: World Bank Institute.
- Kaufmann, D. and A. Kraay. 2002. "Growth without Governance." *Economía* 3(1).
- Knack, S. 2006. "Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators." World Bank Policy Research Working Paper 3968.
- Knack, S. and P. Keefer. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics* 7: 207-27.
- Meisel, N. 2004. *Governance Culture and Development: A Different Perspective on Corporate Governance*. Paris: Development Centre, OECD.
- Metagora. Inventory of Initiatives Aimed at Measuring Human Rights and Democratic Governance. Available online at [http://www.metagora.org/html/aboutus/about\\_inventory.html](http://www.metagora.org/html/aboutus/about_inventory.html) (accessed June 2007).
- North, D.C. 1990. *Institutions, Institutional Change and Economic Performance*.

- Cambridge: Cambridge University Press.
- North, D.C. 2005. *Understanding the Process of Economic Change*. New Jersey: Princeton University Press.
- Oman, C. and G. Wignaraja. 1991. *The Postwar Evolution of Development Thinking*. London: Palgrave MacMillan Ltd. and Development Centre Study.
- PEFA secretariat. 2005. "Public Financial Management Performance Measurement Framework." Washington D.C.: World Bank.
- Razafindrakoto, M. and Roubaud, F. 2006. "Are International Databases on Corruption Reliable? A Comparison of Expert Opinion Surveys and Household Surveys in Sub-Saharan Africa." Paris: DIAL.
- Transparency International. 2004. Frequently Asked Questions about the TI Corruption Perceptions Index (CPI) 2006. Available online at [http://www.transparency.org/policy\\_research/surveys\\_indices/cpi/2006/faq#using](http://www.transparency.org/policy_research/surveys_indices/cpi/2006/faq#using) 2 (accessed June 2007).
- Thomas, M.A. 2007. "What do the Worldwide Governance Indicators measure?" Working Paper, Johns Hopkins University.
- UNDP. 2007. Governance Indicators: A Users' Guide. Available online at [http://www.undp.org/oslocentre/docs07/undp\\_users\\_guide\\_online\\_version.pdf](http://www.undp.org/oslocentre/docs07/undp_users_guide_online_version.pdf). (accessed June 2007).
- Van de Walle, S. 2007. "Contested Definitions of Performance: Evaluating the Performance of Services of General Interest."
- Williamson, J. 2000. "What Should the World Bank Think About the Washington,

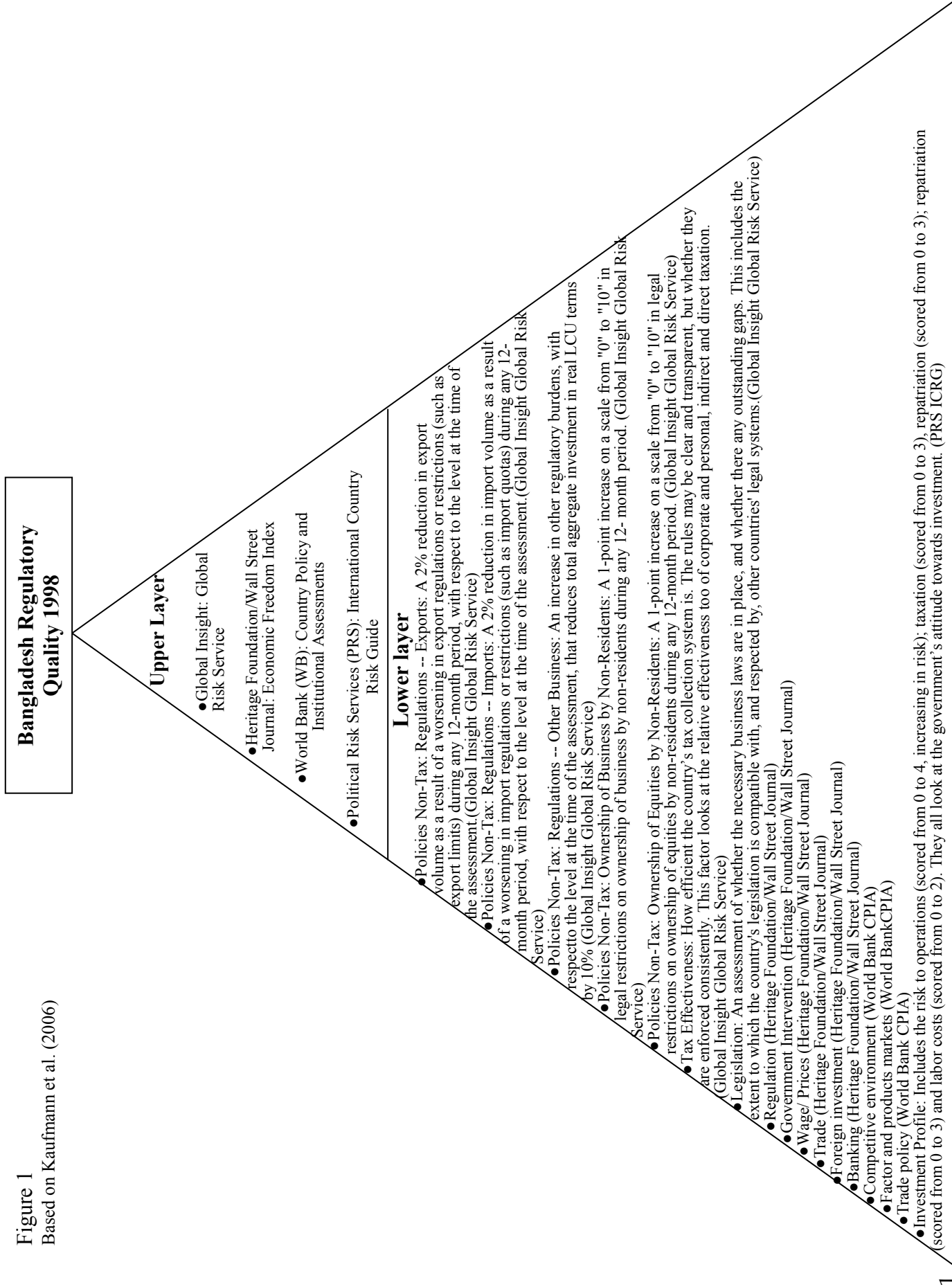
Consensus?" *World Bank Research Observer*

World Bank. 2004. "World Bank Annual Report 2004." Washington, D.C.

World Bank Institute. Governance Data: Web-Interactive Inventory of Datasets and Empirical Tools. Available online at [www.worldbank.org/wbi/governance/govdatasets/index.html](http://www.worldbank.org/wbi/governance/govdatasets/index.html). (accessed June 2007).

World Economic Forum. 2006. *Global Competitiveness Report 2006-2007*: Palgrave Macmillan.

Figure 1  
Based on Kaufmann et al. (2006)



## Bangladesh Regulatory Quality 2005

### Upper Layer

- Asian Development Bank (ADB): Country Policy and Institutional Assessments (CPIA)
- Bertelsmann Foundation: Bertelsmann Transformation Index
- Global Insight: Global Risk Service
- Economist Intelligence Unit (EIU): Country Risk Service
- World Economic Forum (WEF): Global Competitiveness Report, Executive Opinion Survey
- Heritage Foundation/Wall Street Journal: Economic Freedom Index
- Merchant International Group: Grey Area Dynamics
- World Bank (WB): Country Policy and Institutional Assessments
- Political Risk Services (PRS): International Country Risk Guide
- Global Insight: Business Conditions and Risk Indicators

### Lower layer

- Trade Policy (ADB) ● Competitive environment (ADB) ● Factor and products markets (ADB)
- Price stability (Bertelsmann Foundation) ● Private Property (Bertelsmann Foundation) ● Policies: Non-Tax: Regulations -- Exports: A 2% reduction in export volume as a result of a worsening in export regulations or restrictions (Global Risk Service) ● Policies Non-Tax: Regulations -- Imports: A 2% reduction in import volume as a result of a worsening in import regulations or restrictions (such as import quotas) during any 12-month period, with respect to the level at the time of the assessment. (Global Insight Global Risk Service) ● Policies Non-Tax: Regulations -- Other Business: An increase in other regulatory burdens, with respect to the level at the time of the assessment, that reduces total aggregate investment in real LCU terms by 10% (Global Insight Global Risk Service) ● Policies Non-Tax: Ownership of Business by Non-Residents: A 1-point increase on a scale from "0" to "10" in legal restrictions on ownership of business by non-residents during any 12-month period. (Global Insight Global Risk Service) ● Policies Non-Tax: Ownership of Equities by Non-Residents: A 1-point increase on a scale from "0" to "10" in legal restrictions on ownership of equities by non-residents during any 12-month period. (Global Insight Global Risk Service) ● Tax Effectiveness: How efficient the country's tax collection system is. The rules may be clear and transparent, but whether they are enforced consistently. This factor looks at the relative effectiveness too of corporate and personal, indirect and direct taxation. (Global Insight Global Risk Service) ● Legislation: An assessment of whether the necessary business laws are in place, and whether there are outstanding gaps. This includes the extent to which the country's legislation is compatible with, and respected by, other countries' legal systems. (Global Insight Global Risk Service) ● Unfair competitive practices (EIU) ● Price controls (EIU) ● Discriminatory tariffs (EIU) ● Excessive protections (EIU) ● Discriminatory taxes (EIU) ● Administrative regulations are burdensome (WEF) ● Import barriers as obstacle to growth (WEF) ● Competition in local market is limited (WEF) ● It is easy to start company (WEF) ● Anti monopoly policy is lax and ineffective (WEF) ● Environmental regulations hurt competitiveness (WEF) ● Cost of tariffs imposed on business (WEF) ● Government subsidies keep uncompetitive industries alive artificially (WEF) ● Complexity of Tax System (WEF) ● Domestic banks are protected from foreign competition (WEF) ● Barriers to entry in banking sector are very high (WEF) ● Interest rates are heavily regulated (WEF) ● Private sector participation in infrastructure projects is not permitted (WEF) ● Costs of uncertain rules, laws, or government policies (WEF) ● Transfer costs associated with exporting capital as an obstacle to business (WEF) ● General uncertainty on costs of regulations as an obstacle to business (WEF) ● Openness of public sector contracts to foreign investors (WEF) ● Policies for dividend remittances as obstacles to development (WEF) ● Dominance of state owned or state controlled enterprise (WEF) ● State interference in private business (WEF) ● Regulatory discretionality (WEF) Price controls as an obstacle to business development (WEF) ● Regulations on foreign trade as an obstacle to business development (WEF) ● Foreign currency regulations as an obstacle to business development (WEF) ● Regulation (Heritage Foundation/Wall Street Journal) ● Wage/ Prices (Heritage Foundation/Wall Street Journal) ● Trade (Heritage Foundation/Wall Street Journal)
- Foreign investment (Heritage Foundation/Wall Street Journal) ● Banking (Heritage Foundation/Wall Street Journal) ● Unfair Trade (Merchant International Group) ● Unfair Competition (Merchant International Group)
- Competitive environment (World Bank CPIA) ● Factor and products markets (World Bank CPIA) ● Trade policy (World Bank CPIA) ● Investment Profile: Includes the risk to operations (scored from 0 to 4, increasing in risk); taxation (scored from 0 to 3); repatriation (scored from 0 to 2); and labor costs (scored from 0 to 2). They all look at the government's attitude towards investment. (PRS ICRG)
- Tax Effectiveness: How efficient the country's tax collection system is. The rules may be clear and transparent, but whether they are enforced consistently. This factor looks at the relative effectiveness too of corporate and personal, indirect and direct taxation. (Global Insight: Business Conditions and Risk Indicators) ● Legislation: An assessment of whether the necessary business laws are in place, and whether there are any outstanding gaps. This includes the extent to which the country's legislation is compatible with, and respected by, other countries' legal systems. (Global Insight: Business Conditions and Risk Indicators)

- 
- 1 See, for example, Williamson (2000). Analysis of the experiences of the transition economies of the former Soviet Union and Central and Eastern Europe has further strengthened this understanding (see for example, Cornia and Popov, eds., 2001).
- 2 For a detailed description, classification and history of the indicators, see Arndt and Oman (2006). See Hood et al. (2007) for a description and history of the growth in public sector performance rankings
- 3 To illustrate, the World Bank Governance diagnostic surveys in Indonesia in 2001 helped to identify those governance problems that citizens in the country thought to be the most urgent to tackle. This gave an incentive for the government to reform in this area in order to increase their popularity, according to a World Bank staff member. The steering committee consisted mainly of local stakeholders who were also in charge of communicating the results to the public.
- 4 We thank Aart Kraay for this information
- 5 In addition to each country's rating, Kaufmann *et al.* provide so-called "margins of error" to estimate the reliability of the rating. They construct 90 % confidence intervals on the basis of these margins of error around a country's rating which, according to their estimate, has a 90 per cent probability of containing the country's "true" (as opposed to estimated) score. Kaufmann *et al.* warn users not to compare countries among each other or over time where the confidence intervals overlap. To illustrate, users should not conclude that Bangladesh's Regulatory Quality changed from 1998 to 2005, because the respective confidence intervals overlap.
- 6 Several sources are commercial and access is expensive. To illustrate, access to the source Global Risk Data from Global Insight, costs \$ 12 700
- 7 An example of a Global Insight indicator is "How much of a threat businesses face from crime such as kidnapping, extortion, street violence, burglary and so on. These problems can cause major inconvenience for foreign investors and require them to take expensive security precautions".
- 8 This method should not be confused with factor analysis. See Arndt and Oman (2006) for details.
- 9 To illustrate, Guatemala's rating on Regulatory Quality worsened considerably from 1998 to 2005. The rating for 1998 was based on three sources and the rating for 2005 on eight sources, among them the three original sources from 1998. The rating of one of these three sources, the CPIA, is not publicly available for 1998. The other two sources that are available for both years actually rated the country better in 2005 than in 1998. The deterioration in Guatemala's rating from 1998 to 2005 was, hence, very likely driven by the additional sources for the 2005 rating (example based on data from Kaufmann et al., 2006)
- 10 Johnson and Zajonc (2006) refer to the indicators used for the MCC (among them several perception-based indicators like the WIGIs and Freedom House's Political Rights and Civil Liberties indicators) as "a set of objective and transparent governance indicators".
- 11 The Inter American Development Bank (IADB, 2007) lists more than 400 indicators in its database