

MEASURING GOVERNANCE

by
Charles P. Oman and Christiane Arndt

- Quantitative indicators of the quality of governance in developing countries and emerging economies have greatly proliferated since the mid-1990s.
- The main users of these indicators are international investors, official development agencies, journalists and academics.
- The most widely used, and misused, governance indicators are composite perceptions-based indicators.
- Even the most carefully constructed composite indicators have limitations their users seem widely to ignore.
- Greater transparency is required both in the production and in the use of governance indicators.

POLICY BRIEF No. 39

Measuring Governance

by

Charles P. Oman and Christiane Arndt



THE OPINIONS EXPRESSED AND ARGUMENTS EMPLOYED IN THIS PUBLICATION ARE THE SOLE RESPONSIBILITY OF THE AUTHORS AND DO NOT NECESSARILY REFLECT THOSE OF THE OECD, ITS DEVELOPMENT CENTRE OR OF THE GOVERNMENTS OF THEIR MEMBER COUNTRIES.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 33 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

The Development Centre of the Organisation for Economic Co-operation and Development was established by decision of the OECD Council on 23 October 1962 and comprises 25 member countries of the OECD: Austria, Belgium, Chile, the Czech Republic, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Korea, Luxembourg, Mexico, the Netherlands, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey and the United Kingdom. In addition, the following non-OECD countries are members of the Development Centre: Brazil (since March 1994); India (February 2001); Romania (October 2004); Thailand (March 2005); South Africa (May 2006); Egypt and Viet Nam (March 2008); Colombia (July 2008); Indonesia (February 2009); Costa Rica, Mauritius, Morocco and Peru (March 2009) and the Dominican Republic (November 2009). The Commission of the European Communities also takes part in the Centre's Governing Board.

The Development Centre, whose membership is open to both OECD and non-OECD countries, occupies a unique place within the OECD and in the international community. Members finance the Centre and serve on its Governing Board, which sets the biennial work programme and oversees its implementation.

The Centre links OECD members with developing and emerging economies and fosters debate and discussion to seek creative policy solutions to emerging global issues and development challenges. Participants in Centre events are invited in their personal capacity.

A small core of staff works with experts and institutions from the OECD and partner countries to fulfil the Centre's work programme. The results are discussed in informal expert and policy dialogue meetings, and are published in a range of high-quality products for the research and policy communities. The Centre's *Study Series* presents in-depth analyses of major development issues. *Policy Briefs* and *Policy Insights* summarise major conclusions for policy makers; *Working Papers* deal with the more technical aspects of the Centre's work.

For an overview of the Centre's activities, please see www.oecd.org/dev

Also available in French under the title:
La mesure de la gouvernance

OECD 2010

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translations rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.

Table of Contents

Introduction.....	4
Why all the Interest in Governance?.....	5
Deconstructing an Indicator	7
Users Beware! Four dangers	11
Conclusions and Recommendations	16
Notes	20
Annex.....	22
References	24
Other Titles in the Series	25

Introduction

The use of governance indicators, as applied to developing countries, has grown spectacularly in recent years. Following the maxim that you cannot manage what you cannot measure, international investors and official development aid agencies, together with academics and the media, have turned widely to using quantitative governance indicators for both analytical and decision-making purposes – with far-reaching consequences for developing countries.

The most widely used indicators, among the hundreds that have appeared in response to this demand, are composite (or “aggregate”) perceptions-based indicators. Such indicators aggregate often large amounts of information from diverse sources and reduce it to a single number – a single governance score – per country, per year, to facilitate comparisons. The aggregated information consists of people’s perceptions of the quality of governance, or some aspect of governance (e.g., the rule of law, control of corruption), in different countries. Most of the people whose perceptions are used are “experts” or business managers, many of whom live outside the countries they are rating; a few indicators also include local perceptions obtained through household surveys.

Even the most carefully constructed composite indicators suffer from limits that their users seem widely to ignore. These limits notably include the invisible nature of biases built into the indicators, the indicators’ incapacity to guide the efforts of people seeking to improve governance, and the dangers of using the indicators to compare the quality of governance between countries and/or over time.

This *Brief* seeks to clarify the limits of these governance indicators from a user’s perspective – be the user an official aid agency, international investor, researcher, policy analyst or journalist. It seeks above all to reduce the extent to which these indicators are misused.

While the perfect governance indicator does not exist, it is particularly important to increase the transparency of governance indicators, in terms of both their construction and their use. Only through greater transparency can governance indicators contribute more to enhancing the quality of governance in developing countries.

Why all the Interest in Governance?

First, a brief look at the principal causes of the spectacular growth of interest in measuring the quality of governance in developing countries. Four causes appear particularly important.

International Investment

One is the spectacular increase in international investment flows to developing and “emerging” economies. Foreign direct investment into these economies jumped from an annual average of about \$10 billion in the early 1980s to over \$100 billion in the mid-1990s and over \$200 billion after 2004. Even more spectacular has been the growth of international *portfolio* investment – notably by OECD-based institutional investors – whose net annual flows to emerging economies (equity and bond purchases combined) rose from less than \$2 *million* in the late 1980s to over \$50 billion in the 1990s and over \$80 billion in 2000-07.

The spectacular growth of international (largely OECD-based) investors’ assets exposed to risk in developing and emerging economies thus goes far to explain the rapid growth of investors’ interest in the quality of governance in those countries. This interest was further amplified, moreover, by the sea change in economic policy orientation across the developing world. Many developing-country governments became less interventionist, more market-oriented and moved to implement more investor-friendly policy regimes. Driven by growing competition among developing countries to attract foreign investment, this sea change led to a high degree of convergence or homogenisation, at least on paper, of policy regimes across developing countries.

For many international investors, differences between potential host countries’ nominal policy regimes have thus become less important than differences between those countries’ perceived credibility in policy implementation. And the perceived quality of local governance (both political and corporate governance) strongly affects investors’ perceptions of a country’s policy credibility (Oman, 2000). Perceived differences between countries in terms of the quality of local governance have thus emerged as a key determinant of international investment flows to developing and emerging economies¹.

End of the Cold War

The end of the Cold War also contributed to the explosive growth of attention to the quality of governance in developing countries. Throughout the post-war period, OECD governments and their national and multilateral aid agencies sought to promote economic and social development in developing countries, not only as means to fight poverty and raise local living standards, but also to reduce local temptations to turn to communism. OECD governments and aid agencies, both bilateral and multilateral, thus focused on trying to help developing-country governments improve their policies and their people's welfare without significantly questioning the quality of local governance. To challenge pro-western but corrupt or undemocratic aid recipients on their problems of corruption or bad governance was thought likely to prove counter-productive for donors' Cold War objectives.

Only after the demise of the Soviet Union, and the perceived disappearance of the communist threat in the 1990s, did these attitudes and behaviour become susceptible to real change. A watershed was World Bank President James Wolfensohn's 1996 decision to reverse the Bank's longstanding policy of "staying out of politics" and of not officially recognising or seeking to address the acute problems of corruption in many of its borrowing countries. Many other official development agencies, both bilateral and multilateral, then began to give corruption and bad governance in developing countries the attention they deserve as major obstacles to economic development. The Bank's own lending for programmes to improve public governance in its borrowing countries rose to 25 per cent of its total lending by 2004, and continues to grow².

Failed Policy Reform

Also important has been growing recognition of the relative failure of policy reforms promoted through conditional lending by multilateral financial organisations and widely implemented by developing countries in the 1980s and 1990s. Those policy reforms, referred to by some as the "Washington consensus", were reflected in the sea change in economic policy orientation noted earlier.

Beyond the ensuing debate over whether the relative failure of those policy reforms is better explained by too much or rather too little effective implementation of the recommended reforms, what is important for our purposes is the generalised recognition of their relative failure.

This recognition has in turn contributed to a growing understanding – including within the multilateral organisations and among staunch defenders of the importance of market-friendly policy regimes – that strong markets require good governance, and that poor local governance is a likely cause of the relative failure of many of the liberalising policy reforms undertaken in the 1980s and 1990s³. Poor local governance is also perceived as a major potential obstacle to the efficacy of political reforms, and to the success of international cooperation.

New Institutional Economics

A fourth factor, whose importance is great though difficult to measure, is the work of Douglass North and the New Institutional Economics of which he is a leading figure. That work has convincingly demonstrated the importance of a country's system of governance – its formal and informal institutions (the latter including its culture and unwritten values) and their interaction with the behaviour of economic and political entrepreneurs and organisations – for the country's success in terms of its long-term economic growth, enhancement of human welfare and societal development (North, 1990).

Deconstructing an Indicator

The most widely used governance indicators, as noted earlier, are composite perceptions-based indicators. Among these, Transparency International's annually published *Corruption Perceptions Index* (CPI) and the six *Worldwide Governance Indicators* (WGI) published by the World Bank Institute are the most widely used. A key feature of these indicators is their inclusion of an estimated "confidence interval" together with the point score for each country covered by the indicator. Their authors explain, moreover, that differences between countries' point scores whose confidence intervals overlap should be considered statistically insignificant⁴.

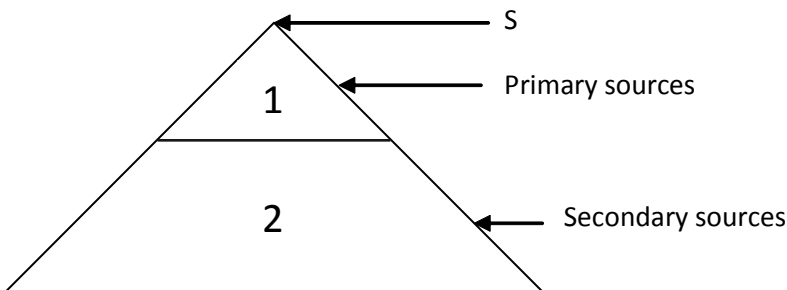
Composite indicators are particularly attractive to users because they reduce complex governance realities to a single number (a "point score") for each country in a given year (a score which is accompanied by an estimate of the score's degree of precision). Composite indicators thus offer a quantitative estimate of the quality of a country's governance, or

particular aspect of its governance, and make it easy for users to compare scores among countries and/or across time.

The fact that many users – in practice if not intention – significantly over-estimate the precision of countries’ governance scores is nevertheless a cause for concern. Part of the problem is that the size of the “confidence intervals” tends to be underestimated (i.e., the point scores’ precision tends to be over-estimated) and may indeed be meaningless, from a statistical perspective, as Anderson (2009) argues. But, above all, the problem is the high degree to which users ignore warnings about the imprecision of scores: users tend widely to use countries’ governance scores as if they were accurate to a degree they are not.

Pyramidal Construction

A good way to visualise both the causes and the consequences of the problem is to think of the Worldwide Governance Indicators (WGI) or the Corruption Perceptions Index (CPI) as a set of pyramids, in which each pyramid corresponds to a country covered by the indicator, where the tip of each pyramid corresponds to a country’s point score in a given year. Graphically, the point “S” in the diagram corresponds to the point score for a country in a particular year.



The sources of information used directly to calculate the score constitute the primary sources and correspond to the upper level (level “1”) in the figure. Each of the numbers in the upper level of the pyramid is itself the result of an aggregation of information and quantitative indicators, often numerous, which can be thought of together as constituting the base, or level “2”, of the pyramid. Annex I provides the pyramids for Bangladesh corresponding to the WGI “Regulatory Quality” indicator for 1998 and 2005.

This way of visualising the often massive amount of information that is reduced to a single, composite number for each country in a given year can be used to draw attention to two other key features of composite indicators.

One is the fact that, for a given indicator in a given year, each country has its own pyramid, and each pyramid is different from all the others because the secondary sources (and often even the primary sources) used to construct the actual score differ from one country to another, for a given indicator in a given year.

The other key feature is that the contents of each pyramid (i.e. the primary and secondary sources of information from which the score of a country is calculated) constantly change over time – they change whether or not the reality of governance in the country, which they are supposed to reflect, changes.

A country's score will thus change from year to year independently of the reality of governance in that country, for two reasons. First, the secondary and therefore the primary sources of information used to calculate the country's score constantly evolve, and do so in ways of which even the producers of the composite indicators themselves cannot be fully aware, at least as regards the secondary sources. Each year, new sources are included and others disappear or are excluded or modified.

Second, perceptions of the quality of governance in a country – and it is perceptions that determine the scores – can and do easily change in the absence, or independently, of any change in the objective reality of governance in the country.

The result is that nobody – neither the users nor even the producers of a perceptions-based composite indicator – can know all the changes occurring down inside a pyramid that explain the change (or lack of change) in a country's governance score from one year to another.

The Example of Regulatory Quality in Bangladesh

The Worldwide Governance Indicator on "Regulatory Quality" for Bangladesh, provided in Annex I for 1998 and 2005, provides a good illustration⁵. We should also note that each year the scores of the 200-odd countries covered by a WGI are normalised, so that all countries' scores fall between -2.5 and +2.5.

In 1998, Bangladesh's score was -0.13, which put it in the 39th percentile (starting from the lowest, or worst) of the approximately 200 country scores (Kaufmann, *et al.*, 2006). That score was derived from four primary sources, and the estimated "90 per cent confidence interval" spanned the scores between the 19th and the 66th percentiles; the score's estimated degree of precision was quite low, in other words, because the interval was quite large. This low degree of precision was undoubtedly due at least in part to the relatively low number of primary sources used to produce the score.

In 2005, Bangladesh's "Regulatory Quality" WGI score was -1.07, which put it into the 15th percentile of the 200-odd countries covered by this indicator that year. The score was derived from ten primary sources, and the estimated "90 per cent confidence interval" was between the 8th and the 22nd percentiles. The score's estimated degree of precision was thus higher in 2005 than in 1998 – due, at least in part, to the larger number of primary sources used to produce it.

This apparent decline in Bangladesh's regulatory quality between 1998 and 2005 – reflected in the fall in the point score from -0.13 to -1.07, corresponding to a fall from the 39th to the 15th percentile – can be explained by one or more of four possible causes:

(1) Regulatory quality actually declined in Bangladesh between 1998 and 2005. (It is nevertheless interesting to note that, during this same period, the country witnessed relatively strong economic growth.)

(2) Perceptions of regulatory quality in Bangladesh deteriorated (independently of any objective change, or lack of change, in Bangladesh's regulatory quality) during this period.

(3) Perceptions of regulatory quality in other countries covered by this indicator improved (independently of any change or lack of change of regulatory quality in Bangladesh) during this period.

(4) The change in Bangladesh's score between 1998 and 2005 is a statistical artifice caused by a modification in the sources of information, down inside the pyramid, used to construct the scores in the two different years.

All four of these possible explanations may have contributed to the fall in Bangladesh's WGI "Regulatory Quality" score from the 39th to the 15th percentile between 1998 and 2005. The fourth possible explanation nevertheless deserves particular attention, because of the significant increase in the number of primary sources used to calculate the score –

from four in 1998 to ten in 2005. This increase goes far to explain the significant reduction in the estimated size of the confidence interval, and probably also goes far to explain the change in the point score.

This example also usefully reminds us, as the producers of the WGI would themselves have to agree, that the overlap of the confidence intervals for the 1998 and 2005 scores means that any difference between those two scores should be interpreted as statistically insignificant – i.e. that the apparent decline in Bangladesh’s regulatory quality is statistically meaningless. The authors’ own “rule of thumb”, according to which any apparent difference between point scores whose confidence intervals overlap should be interpreted as statistically insignificant, is important to underline, because it is too often ignored in practice by users of those indicators.

Users Beware! Four dangers:

Users should be aware of four mutually reinforcing problems associated with the use of these indicators, largely inherent in the very nature of the indicators:

Transparency Paradox

The first of these problems, which aggravates the others as well, is the considerable lack of transparency of the indicators’ content. This problem is not only important, but also paradoxical, because of the considerable extent to which official aid agencies use these indicators to try to enhance the transparency and objectivity of their aid-allocation decisions.

Why are the most widely used governance indicators much less transparent than many of their users seem to realise? Three reasons stand out:

(A) One is the sheer complexity and enormous amount of subjective information the indicators reduce to a single number – and do so using sources of information whose weights vary from country to country, and change all the time even for a given country.

(B) Another key factor limiting the indicators’ effective transparency is the fact that not all the information used to calculate the indicators is

publicly available, which makes it impossible for others to replicate or reconstruct the indicators and get the same results⁶.

(C) A third, and in key respects ultimately the most important factor limiting the indicators' transparency, is their lack of any underlying conceptual or analytical framework, or theory of governance, to inform or guide their scoring systems. Countries' scores thus do not reflect any theory or concept of "good" (or "bad") governance. They only reflect the procedure used to aggregate and give relative weights to the perceptions chosen for inclusion in the indicator.

It is nevertheless important to remember that all ratings systems are, by their very nature, normative. They imply a judgement. One would therefore expect to have a normative definition or concept of what constitutes "good" or "not good" or "less good" governance which informs or guides the process that determines countries' ratings. But there is no such normative concept or theory underlying these ratings.

The WGI's producers define governance simply as "the traditions and institutions by which authority in a country is exercised." They portray the six WGI indicators as comprising two dimensions in each of three distinct channels by which "authority is exercised", with each of the six indicators thus portrayed as capturing a specific dimension of "governance"⁷.

While the statistical validity of this disaggregation has been sharply questioned (Langbein and Knack, 2008), users need above all to understand that the producers of the WGI build on no normative theory, criteria or analytical framework that explains or gives guidance on how to judge or evaluate the quality of "the traditions and institutions by which authority in a country is exercised." They do not because they cannot. They cannot, because it is the aggregation of subjective perceptions – the data down inside each pyramid, together with the way those data are aggregated – which gives *de facto* content or "meaning" to the normative definition of "good" or "not good" (or "less good") governance for each pyramid, i.e., for each country in a given year. And, again, not only is every pyramid different, to varying degrees, from every other pyramid in a given year; the pyramid for a given country changes over time. All the pyramids change, all the time.

An important consequence is that nobody – including the producers of the WGI themselves – can say what those norms are. The norms are in fact determined, implicitly though not explicitly, for each country in a given year, by the subjective views of all those people whose perceptions are

aggregated into that specific pyramid, and by the aggregation techniques used to reduce those views to a single number. They are not determined by any definition, concept or theory of what constitutes *good* governance, or of how to distinguish good from bad governance.

Hidden Bias

The second problem, exacerbated by the indicators' lack of transparency, is the hidden bias built into the indicators. The issue is not bias as such – all normative instruments will have some bias – but the fact that whatever biases the indicators may embody remain largely hidden from the indicators' users.

A good illustration of such hidden bias is the way regulations to protect workers' rights or the environment tend to be interpreted as market-unfriendly by primary sources that carry the most weight in the aggregation process. Users may or may not share the value judgement embodied in this process. Our concern is that most users simply will not be aware of this type of bias built into the indicators.

Another good illustration is the very low effective weight given to household surveys in the aggregation process, relative to the weights given to firm surveys and "expert" assessments. One reason why household surveys are given little weight is that such surveys are relatively expensive to carry out, and among the surveys that are made, very few cover households in more than a single country – whereas only data sources that cover more than a single country can be used for cross-country comparative indicators such as these governance indicators.

The bias away from the perceptions reflected in household surveys, compared to the weight given to perceptions of "experts" and firm surveys, is thus caused in part by a relative paucity of household survey data. But the bias is made much stronger by the fact that the creators of the WGI have chosen to use an aggregation technique that gives less weight to perceptions that diverge from the dominant majority of views. Yet household surveys tend to capture the views of the poor much better than "expert" assessments and firm surveys. Insofar as the views of the poor differ from those of firm managers and "experts" they will be given even less weight in the construction of the indicators, to the point where in practice they carry very little weight⁹. Thus, for example, the WGI "rule of law" indicator gives zero weight to Gallup's worldwide poll of local citizens' perceptions of their exposure to crime, whereas it gives

the third highest weight to Global Insight's business risk and conditions, a Boston-based provider of commercial information that estimates the crime risk to businesses.

Our point, again, is not that the perceptions reflected in household surveys should necessarily be given more weight in the production of governance indicators – although many would argue they should – for this ultimately depends on a given user's specific needs. Our point is rather that in simply listing household surveys, alongside firm surveys and expert assessments, as included among the sources of data embodied in their governance indicators, the WGI's producers give users no indication or warning that the views captured by household surveys actually carry negligible weight in the final scores, and that those scores are very largely determined by the views of "experts" and firm managers. While one could argue that the indicators are effectively biased against the views expressed in household surveys, our deeper concern is not one of bias but of hidden bias.

Not Useful to Guide Concrete Actions

The third problem, after lack of transparency and hidden bias, is that the indicators give little guidance to local stakeholders, or others, as to what concrete action they can or should take to actually improve the quality of local governance.

Thus, for example, a "rule of law" indicator may provide valuable information on how secure business people feel about their property, which can in turn be useful for predicting and perhaps even improving investment behaviour in the country. But the indicator says little about what specifically makes people feel more or less secure about their property in a given country. It thus gives little guidance about how concretely to improve the rule of law or make people feel more secure about their property in any given country.

Governance indicators are used mainly as a decision tool for outsiders who wish to judge or compare countries. But they are not action-oriented for stakeholders in the countries that are being judged or compared. In the jargon of the literature, they are not "actionable".

We are not blaming the producers of the indicators for this limitation. It is inherent in the very nature of the indicators. What is important is that users – donors of development assistance, for example – be aware of it.

Exaggerated Expectations

The fourth problem, and in many ways the one that concerns us most of all, is that users seem widely to over-estimate the degree of accuracy, and therefore the reliability, of countries' point scores for cross-country comparisons and, especially, for identifying change (or lack of change) in the quality of governance over time.

Users understandably seek to compare countries and assess changes in the quality of governance over time. Yet both our conceptual understanding and our ability to measure the quality of governance are in their infancy. Summarising available information on perceptions of governance, which is the purpose of composite indicators like the WGI, can be useful, but if misused can also mislead and/or delay progress. Users must be careful not to simply assume that such indicators match their expectations.

Moreover, while we appreciate efforts to estimate the degree of imprecision of point scores through the estimation of "90 per cent confidence intervals" in the WGI and CPI – if only because they help draw attention to the existence of such imprecision – we also worry about the apparent bias in the estimation of those intervals. We are concerned because the construction of the confidence intervals is based on an assumption that margins of error are not correlated among the different sources of information used to calculate the point scores (thereby also assuming that a greater number of sources increases the degree of precision of the point score). The problem is that considerable empirical evidence suggests, on the contrary, that different sources of information used to calculate the point scores *do* in fact tend to correlate with each other (Arndt and Oman, 2006; Knack, 2007; Arndt, 2009).

The argument that the published "confidence intervals" are not strictly meaningful statistically (Anderson, 2009) does not preclude the possibility that their publication nevertheless serves the useful purpose of helping to draw users' attention to the point scores' imprecision. But the built-in bias of those intervals towards over-estimating the precision of point scores means that their publication may in fact have the opposite effect: by biasing upwards the implied accuracy of point scores, "confidence intervals" may give users greater confidence in point scores' accuracy than is warranted. The result may therefore actually be to aggravate the problem of users' widespread failure to take account of point scores' imprecision.

The net result, in sum, is that comparing the quality of governance across countries and especially over time is even more problematic than we suggested previously.

Conclusions and Recommendations

Measuring governance remains problematic. The lack of a reasonably credible theory to explain causal relationships between specific governance features in a country and the process of development in that country – a theory that could provide the analytical framework within which to define reliable facts-based governance indicators – explains, in part, the widespread use of composite perceptions-based governance indicators.

The principal attraction of these widely used indicators is that they reduce to a single number, for a given country in a given year, complex and often poorly understood governance realities. They do so for many countries, moreover, in a way that makes those numbers look comparable. The result is to give users the impression that they can safely use these indicators to compare the quality of governance across countries and over time.

This impression is reinforced by the well documented tendency of people to believe that numbers are facts. Careless or misleading advertising (e.g. the World Bank Institute has advertised the WGI as “reliable measurements of governance”) has also undoubtedly helped to reinforce this impression⁹.

The principal problem with the way many users employ these indicators is not the subjective nature of the information used to construct them, i.e. that the indicators are perceptions-based. The lack of a theory able to provide an analytical framework within which to define meaningful facts-based governance indicators makes perceptions-based indicators useful and, ultimately, no more subjective in the results they generate than facts-based indicators¹⁰.

Rather, the principal problems – and our most serious concerns – are the hidden biases embodied in the construction of these indicators and, above all, users’ widespread ignorance or under-estimation of the limits to the precision of the most widely used governance indicators. This ignorance or under-estimation is manifest and widespread in the uses

made of these indicators by official aid agencies, international investors, policy analysts and even academic researchers.

Our first recommendation, therefore, is to avoid using complex composite indicators such as the WGI for resource allocation and certainly for any automatic decision-making purposes¹¹. Instead we encourage resource allocation decision makers, notably aid agencies, to rely on more specific indicators whose sources, and the weights attributed to each source, are more transparent and theoretically founded.

Various handbooks and databases can help users find transparent indicators for their purposes, and more easily grasp the specific strengths and limitations of particular indicators. The Inter-American Development Bank, for example, provides useful information on how 400 different indicators are constructed, and on the implications of the different methods of construction for the indicators' reliability, validity, and suitability for making comparisons across countries and over time¹². Also very useful are the UNDP handbooks on different types of governance indicators that are freely available¹³.

While there is no single best governance indicator – if only because there is no consensus view of which specific governance features matter most for successful development today – there are many initiatives to produce transparent governance indicators. One good example, produced since 2001, is the Public Expenditure and Financial Accountability (“PEFA”) indicator. PEFA is jointly managed and financed by the World Bank’s Development Grant Facility, the European Commission, the U.K. Department for International Development, the Swiss State Secretariat for Economic Affairs, the Royal Norwegian Ministry of Foreign Affairs, the French Ministry of Foreign Affairs, and the International Monetary Fund¹⁴.

Another example is the Global Integrity Index, which assesses the existence, effectiveness, and citizens’ access to key national anti-corruption mechanisms used to hold governments accountable. All the underlying data and criteria used to rank the 55 countries it covers are freely accessible to the public¹⁵.

A third example is the OECD’s bi-annual publication *Government at a Glance*¹⁶ whose coverage will be extended to ten non-OECD countries including Brazil and China. The publication compares countries’ political and institutional government frameworks, their public revenues, expenditures and employment, and their government policies and practices in terms of regulatory management, budgeting, integrity and e-government.

The lack of consensus view on which specific features of governance are most important for successful development – a lack which gives the construction of governance indicators an inherent degree of subjectivity – also means that transparent indicators can provide an excellent basis for debate, and for advancing our understanding of which governance features matter most for developing countries today.

Our recommendation to rely on transparent indicators is even more relevant for users, domestic or foreign, whose objective is not merely to rank or judge but ultimately to help improve the quality of governance in a country. For these users, composite indicators may be of some value for advocacy and/or agenda-setting purposes, but they provide little basis for guiding direct actions to raise the quality of governance.

While users often employ indicators they see others using (herd effects), those seeking guidance for concrete efforts to enhance the quality of governance (“actionability”) require more in-depth and country-specific governance indicators and assessments. Among the growing supply of such indicators and assessments that we can recommend, the World Bank Institute’s country-specific Governance Diagnostic Surveys¹⁷ usefully allow for a triangulation of results to compare perceptions of households, business managers and public officials. Also noteworthy are the “mirror surveys” published by the French public research unit “DIAL” (*Développement, Institutions et Analyses de Long terme*) which compare answers from “experts” and local households to the same questions about their perceptions of corruption¹⁸, and the OECD’s analytical overview of governance assessments conducted by bi- and multilateral donors¹⁹.

For producers of governance indicators, we recommend above all to ensure full disclosure and transparency of sources and methods. Users need to be able easily and fully to understand the relative weights effectively given to the views coming from different sources – notably those derived from household surveys relative to those from local and foreign “experts” and business leaders. Producers should therefore publish the values of all sources used to calculate composite point scores together with the relative weights effectively given to each source.²⁰

Producers who publish estimated confidence intervals together with countries’ point scores should provide a clear explanation of how the confidence intervals are estimated, including all underlying assumptions. They should also clearly remind readers of the imprecision of point scores, as should all producers of composite indicators.

Last but not least, producers should explain, with as much clarity and precision as possible, the link between the indicators they publish and the definition of governance, or of the specific aspect(s) of governance, to which each indicator refers.

Notes

1. This trend among investors to focus on the quality of local governance as a key determinant of their investment-location decisions has been reinforced by their growing mistrust of country-risk indicators since the failure of those indicators to foresee the emerging-market financial crises of the late 1990s.
2. From 2003 to 2008, World Bank lending to improve public sector governance rose by 76 per cent, while that to improve economic management fell by 49 per cent (World Bank, 2008).
3. Poor local governance is seen as the unique or principal cause of that failure in the eyes of some, and simply as an additional cause in the eyes of others.
4. Few producers of indicators warn potential users of the degree of imprecision associated with their indicators. Efforts by the producers of the WGI and CPI to do so are therefore laudable. However, as a recent World Bank paper (Anderson, 2009) explains, the estimated “confidence intervals” they publish do not correspond to the traditional statistical interpretation of “margin of error”, which raises serious questions of how statistically to interpret the meaning, if any, of the “confidence intervals” published together with the point scores of the WGI and CPI.
5. We choose this example simply because one of us prepared it to participate in a seminar on governance indicators in Bangladesh in 2006. We continue, in the remainder of this Brief, to focus on the WGI to illustrate our arguments because they are among the most widely used and most influential *governance* indicators, whereas the CPI focuses on one specific dimension of governance.
6. The cost of purchasing access to the WGI’s commercially provided sources is estimated to be about \$90,000 (Iqbal and Shah, 2008). The World Bank Institute’s recent decision to publish a large majority of primary sources for the WGI constitutes an important step in the right direction. But in the absence of the availability of *all* primary and secondary sources, these indicators still cannot in practice be replicated or reproduced by others.
7. The six WGI are “Voice and Accountability” and “Political Stability”, which the authors portray as constituting the political dimension, “Government Effectiveness” and “Regulatory Quality” the economic dimension, and “Rule of Law and “Control of Corruption” the institutional dimension.
8. By way of illustration, the WGI “rule of law” indicator gives significant weight to an indicator of “crime risk to businesses” supplied by a Boston-based provider whereas Gallup’s worldwide poll of citizens’ exposure to crime gets zero weight. It is therefore not surprising that a survey of local governance stakeholders in Rwanda, including both civil-society and government respondents, found the most widely used governance indicators to be biased and inaccurate (Damon and Rudert, 2008).
9. Transparency International advertises its CPI more accurately and appropriately as “a snapshot of the views of business people and country analysts, with less of a focus on year-to-year trends.”
10. Facts-based governance indicators embody a high degree of subjectivity both because of the inherent subjectivity involved in the way the meaning of a fact or set of facts is *interpreted* in the design of a facts-based indicator, by the indicator’s producer(s), and because of the very limited availability and reliability of data needed to construct such indicators for many developing countries.
11. The World Bank itself thus notes that, “The WGI are not used by the World Bank Group to allocate resources or for any other official purpose.”
12. See www.iadb.org/datagob .
13. See www.undp.org/governance/guidelines-toolkits.htm. See also Arndt and Oman (2006), Knack (2007) and Arndt (2009) for additional guidelines and analytical overviews.

14. See www.pefa.org/assessment_reportmn.php .
15. See www.globalintegrity.org
16. See www.oecd.org/gov/indicators/govataglance
17. See www.worldbank.org/wbi/governance/diagnostics .
18. See Razafindrakoto and Roubaud (2006), available in English and French, at www.dial.prd.fr/dial_publications/PDF/Doc_travail/2006-17_english.pdf .
19. See www.oecd.org/dac/governance/govassessment .
20. The OECD has published in cooperation with the European Commission a handbook on constructing composite indicators which gives step-by-step advice on the construction of transparent and high-quality indicators (cf. www.oecd.org/document/9/0,3343,en_2649_34349_41752777_1_1_1_1,00.html).

Annex

BANGLADESH REGULATORY QUALITY 1998

UPPER LAYER

- Global Insight: Global Risk Service • Heritage Foundation / Wall Street Journal: Economic Freedom Index • World Bank (WB): Country Policy and Institutional Assessments • Political Risk Services (PRS): International Country Risk Guide

LOWER LAYER

- Policies Non-Tax: Regulations – Exports: A 2% reduction in export volume as a result of a worsening in export regulations or restrictions (such as export limits) 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 – 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 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 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 Global Risk Service)
- Regulation (Heritage Foundation/Wall Street Journal) • Government Intervention (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)
- 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 3); repatriation (scored from 0 to 3) and labor costs (scored from 0 to 2). They all look at the government's attitude towards investment. (PRS ICRG)

**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 • 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 (such as export limits) 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 – 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 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 Global Risk Service)
- Unfair competitive practices (EIU) • Price controls (EIU) • Discriminatory tariffs (EIU) • Excessive protections (EIU) • Discriminatory taxes (EIU)
- Administrative regulations are burdensome (WEF)

Source: based on Kaufmann *et al.* (2006).

References

- ANDERSON, J. (2009), "A Review of Governance and Anti-corruption Indicators in East Asia and Pacific", draft working paper, World Bank, Washington, DC, June.
- ARNDT, C. (2009), "Governance Indicators", doctoral dissertation, Maastricht University School of Governance, Maastricht, June.
- ARNDT, C. (2008), "The Politics of Governance Ratings" in *International Public Management Journal*, July
- ARNDT, C. and C. OMAN (2006), *Uses and Abuses of Governance Indicators*, OECD Development Centre Studies, Paris.
- DAMON, M. and L. RUDERT (2008), "International Governance Indices – Making Them Work for Developing Countries", Policy Analysis Exercise for the OECD Development Centre, JFK School of Government, Harvard University, Cambridge, June.
- IQBAL, K. and A. SHAH (2008), "How Do Worldwide Governance Indicators Measure Up?", World Bank working paper, Washington DC, April.
- KAUFMANN, D. *et al.* (2006), "Governance Matters V: Governance Indicators for 1996-2005" World Bank Institute, Washington, DC.
- KNACK, S. (2007), "Measuring Corruption: A Critique of Indicators in Eastern Europe and Central Asia" in *Journal of Public Policy*, December.
- LANGBEIN, L. and S. KNACK (2008), "The Worldwide Governance Indicators and Tautology: Causally Related Separable Concepts, Indicators of a Common Cause, or Both?" World Bank Policy Research Working Paper 4669, World Bank, Washington, DC, July.
- NORTH, D. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge.
- OMAN, C. (2000), *Policy Competition for Foreign Direct Investment: A Study of Competition Among Governments to Attract FDI*, OECD Development Centre, Paris.
- RAZAFINDRAKOTO, M. and F. ROUBAUD (2006), "Are International databases on corruption reliable? A comparison of expert opinion surveys and household surveys in sub-Saharan Africa", *Working Paper 2006*, DIAL, Paris, December.
- WORLD BANK (2008), *Annual Report*, World Bank, Washington, DC.

Other Titles in the Series

The Development Centre *Policy Briefs* can be downloaded from: www.oecd.org/dev/briefs
or obtained by e-mail : dev.contact@oecd.org

Adjustment and Equity (No. 1)

by Christian Morrisson (January 1992)

Managing the Environment in Developing Countries (No. 2)

by David O'Connor and David Turnham (January 1992)

Privatisation in Developing Countries: Reflections on a Panacea (No. 3)

by Olivier Bouin (April 1992)

Towards Capital Account Convertibility (No. 4)

by Bernhard Fischer and Helmut Reisen (April 1992)

Trade Liberalisation: What's at Stake? (No. 5)

by Ian Goldin and Dominique van der Mensbrugge (May 1992)

Towards Sustainable Development in Rural Africa (No. 6)

by David Turnham, with Leif E. Christoffersen and J. Thomas Hexner (January 1993)

Employment Creation and Development Strategy (No. 7)

by David Turnham (July 1993)

The Disarmament Dividend: Challenges for Development Policy (No. 8)

by Jean-Claude Berthélemy, Robert S. McNamara and Somnath Sen (April 1994)

Pension Fund Investment from Ageing to Emerging Markets (No. 9)

by Bernhard Fischer and Helmut Reisen (January 1995)

What Institutional Framework for the Informal Sector? (No. 10)

by Christian Morrisson (October 1995)

The Policy Challenges of Globalisation and Regionalisation (No. 11)

by Charles Oman (June 1996)

Policies for Economic Take-off (No. 12)

by Jean-Claude Berthélemy and Aristomène Varoudakis (September 1996)

The Political Feasibility of Adjustment (No. 13)

by Christian Morrisson (October 1996)

Biotechnology Policy for Developing Country Agriculture (No. 14)
by Carliene Brenner (April 1997)

Pension Reform: Lessons from Latin America (No. 15)
by Monika Queisser (January 1999)

After the Great Asian Slump: Towards a Coherent Approach to Global Capital Flows (No. 16)
by Helmut Reisen (January 1999)

Participatory Governance: The Missing Link for Poverty Reduction (No. 17)
by Hartmut Schneider (April 1999)

Multilateral Tariff Liberalisation and the Developing Countries (No. 18)
by Sébastien Dessus, Kiichiro Fukasaku and Raed Safadi (September 1999)

Health, Education and Poverty Reduction (No. 19)
by Christian Morrisson (October 2001)

The New Regionalism in Sub-Saharan Africa: More Than Meets the Eye? (No. 20)
by Andrea Goldstein (March 2002)

Beyond Johannesburg: Policies and Finance for Climate-Friendly Development (No. 21)
by Georg Caspary and David O'Connor (August 2002)

Strengthening Participation in Public Expenditure Management: Policy Recommendations for Key Stakeholders (No. 22)
by Jeremy Heimans (October 2002)

Corporate Governance in Developing, Transition and Emerging-Market Economies (No. 23)
by Charles Oman, Steven Fries and Willem Buiter (September 2003)

Innovative Approaches to Funding the Millennium Development Goals (No. 24)
by Helmut Reisen (April 2004)

Which Policies Can Reduce the Cost of Capital in Southern Africa? (No. 25)
by Martin Grandes and Nicolas Pinaud (September 2004)

Policy Coherence Towards East Asia: Development Challenges for OECD Countries (No. 26)
by K. Fukasaku, M. Kawai, M.G. Plummer and A. Trzeciak-Duval (April 2005)

Changing Social Institutions to Improve the Status of Women in Developing Countries? (No. 27)
by Johannes Jütting and Christian Morrisson (July 2005)

Harnessing Aid, Trade and Migration for Development: How to Gain from Policy Coherence (No. 28)
by Jeff Dayton-Johnson and Louka T. Katseli (August 2006)

Natural Disaster and Vulnerability (No. 29)
by Jeff Dayton-Johnson (September 2006)

- Migration and Development: Challenges for Policy Making (No. 30)*
by Louka T. Katseli, Robert E.B Lucas and Theodora Xenogiani (October 2006)
- After Gleneagles: What Role for Loans in ODA? (No. 31)*
by Daniel Cohen, Pierre Jacquet and Helmut Reisen (October 2006)
- Commodity Funds: How to Fix Them? (No. 32)*
by Daniel Cohen, Thibault Fally and Sébastien Villemot (November 2006)
- New Actors in Health Financing: Implications for a Donor Darling (No. 33)*
by Denis Drechsler and Felix Zimmermann, (February 2007)
- Banking on Development: Private Banks and Aid Donors in Developing Countries (No. 34)*
by Javier Santiso (February 2008).
- Building Public Awareness of Development Communicators, Educators and Evaluation (No. 35)*
by Annette Scheunpflug and Ida McDonnell (June 2008)
- Making the Most of Aid: Challenges for Africa's Agribusiness (No. 36)*
by Jeff Dayton-Johnson and Kiichiro Fukasaku (July 2008)
- To Benefit from Plenty: Lessons from Chile and Norway (No. 37)*
by Gøril Havro and Javier Santiso (September 2008)
- How to Spend It: Commodity and Non-commodity Sovereign Wealth Funds (No. 38)*
by Helmut Reisen (December 2008)

DEVELOPMENT CENTRE POLICY BRIEFS

In its research activities, the Development Centre aims to identify and analyse problems the implications of which will be of concern in the near future to both member and non-member countries of the OECD. The conclusions represent a contribution to the search for policies to deal with the issues involved.

The *Policy Briefs* deliver the research findings in a concise and accessible way. This series, with its wide, targeted and rapid distribution, is specifically intended for policy and decision makers in the fields concerned.

Official development aid agencies, international investors, academics and the media have greatly increased their use of quantitative indicators to assess the quality of governance in developing countries – with far reaching consequences for these countries. The most widely used indicators suffer from limits and biases which their users often ignore. This *Policy Brief* seeks to clarify these limits and reduce the misuse of governance indicators.

OECD DEVELOPMENT CENTRE
2, rue André-Pascal,
75775 Paris Cedex 16, France
Tel.: + 33 (0)1 45 24 82 00
Fax: + 33 (0)1 44 30 61 49
dev.contact@oecd.org
www.oecd.org/dev/briefs



CENTRE DE DÉVELOPPEMENT
DEVELOPMENT CENTRE

