Development Centre Studies

Uses and Abuses of Governance Indicators

By Christiane Arndt and Charles Oman



OECD
OECDPUBLISHING
DEVELOPMENT CENTRE
CENTRE DE DÉVELOPPEMENT

Development Centre Studies

Uses and Abuses of Governance Indicators

by
Christiane Arndt and Charles Oman

DEVELOPMENT CENTRE OF THE ORGANISATION
FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 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, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, 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.

Also available in french under the title:

Les indicateurs de gouvernance

Usages et abus

© OECD 2006

No reproduction, copy, transmission or translation of this publication may be made without written permission. Applications should be sent to OECD Publishing: *rights@oecd.org* or by fax (33 1) 45 24 13 91. Permission to photocopy a portion of this work should be addressed to the Centre français d'exploitation du droit de copie, 20, rue des Grands-Augustins, 75006 Paris, France (*contact@cfcopies.com*).

THE DEVELOPMENT CENTRE

The Development Centre of the Organisation for Economic Co-operation and Development was established by decision of the OECD Council on 23rd October 1962 and comprises 21 member countries of the OECD: Austria, Belgium, the Czech Republic, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Korea, Luxembourg, Mexico, the Netherlands, Norway, Portugal, Slovak Republic, Spain, Sweden, Switzerland and Turkey, as well as Brazil since March 1994, Chile since November 1998, India since February 2001, Romania since October 2004, Thailand since March 2005 and South Africa since April 2006. The Commission of the European Communities also takes part in the Centre's Governing Board.

The purpose of the Centre is to bring together the knowledge and experience available in member countries of both economic development and the formulation and execution of general economic policies; to adapt such knowledge and experience to the actual needs of countries or regions in the process of development and to put the results at the disposal of the countries by appropriate means.

The Centre is part of the "Development Cluster" at the OECD and enjoys scientific independence in the execution of its task. As part of the Cluster, together with the Centre for Co-operation with Non-Members, the Development Co-operation Directorate, and the Sahel and West Africa Club, the Development Centre can draw upon the experience and knowledge available in the OECD in the development field.

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.

ISBN 92-64-02685-1 © OECD 2006

Foreword

This study is a product of the Development Centre's work on the nature of interactions between the quality of governance and the behaviour of investment in developing countries.

ISBN 92-64-02685-1 © OECD 2006

Acknowledgements

The authors wish above all to thank Denis de Crombrugghe, at the University of Maastricht, for his indefatigable advice and guidance on statistical and econometric issues. They thank Daniel Kaufmann and Aart Kraay at the World Bank Institute for their frank and friendly critical feedback on an earlier version of this study. They thank Marie Wolkers and the team at Transparency International for constructive discussions and their provision of useful material. They thank Orsetta Causa and Ekkehard Ernst (OECD Economics Department), Thomas Heimgartner (Metagora), Javier Herrera, Mireille Razafindrakato and François Roubaud (DIAL), Nico Jaspers and Juan Ramón de Laiglesia (OECD Development Centre), Jana Malinska and Nick Manning (OECD Governance Directorate), Chris de Neubourg (Maastricht Graduate School of Governance), Richard Öhrvall (Statistics Sweden), Jacques Ould-Aoudia and Jan Robert Suesser (France's Ministry of the Economy, Finance and Industry), John Sullivan (CIPE) and participants in the November 2005 informal Development Centre Seminar and the May 2006 Maastricht Graduate School of Governance seminar for their encouragement and comments. They thank Nicolas Meunier (Crédit Agricole) and Helmut Reisen (OECD Development Centre) for sharing their technical expertise on financial markets and country risk analysis. They thank Vanda Legrandgérard for her help in preparing the typescript for publication. All mistakes remain theirs alone.

The authors are also very grateful to the companies, banks and donors of official development assistance interviewed for this study who requested anonymity but whose assistance was invaluable for our work.

The Development Centre acknowledges with gratitude the financial support from the Swiss Agency for Cooperation and Development and from the Washington D.C.-based Center for International Private Enterprise, which made this study possible. It also thanks the Maastricht Graduate School of Governance for its collaboration.

ISBN 92-64-02685-1 © OECD 2006

Table of Contents

Preface *Louka T. Katseli* 9
Summary 11
Introduction 13
Chapter 1 Why all the Interest in Governance? 15
Chapter 2 Sources of Governance Indicators 21
Chapter 3 Uses of Governance Indicators 35
Chapter 4 In-depth Analysis of KKZ Composite Indicators 49
Chapter 5 Governance and Growth 77
Chapter 6 Moving Forward 89
Appendix I Aggregation Methodology for the KKZ Composite Indicators 103
Appendix II Governance and Growth 109
Bibliography 115

ISBN 92-64-02685-1 © OECD 2006

Preface

Governance has come to the fore in recent years as vitally important in developing countries both for international investors and for providers of official development assistance.

For investors, as the OECD's Business and Industry Advisory Committee recently noted, the quality of governance has become the single most important determinant of investment-location decisions in developing and "emerging market" economies. For OECD development co-operation agencies, both national and multilateral, the watershed was in 1996, when the World Bank reversed its long-standing policy of largely ignoring problems of weak institutions and bad governance in borrowing countries; since then, donors have increasingly used governance indicators to identify and reward developing countries that improve the quality of their governance.

Building on its recent findings on the importance of corporate governance in developing and emerging-market economies, the Development Centre has been analysing for some time the nature of interactions between the quality of governance (corporate and public governance included) and investment behaviour. In the process of this work, the limitations of governance indicators used by investors and donors in their investment and aid-allocation processes became evident; these limitations may in fact hamper the usefulness of the indicators for policy makers, investors and donors. Given the importance of appropriate measurement for good management, the purpose of this study is to highlight these limitations with a view towards reducing the misuse of governance indicators and improving their construction and usefulness.

Based on a careful scrutiny of the indicators most widely used today for investment and aid-allocation processes, the analysis has greatly benefited from information gained by the authors through personal interviews with the management and staff of ten major corporate and financial investors, as well

ISBN 92-64-02685-1 © OECD 2006

as with spokespersons for business associations and official development co-operation agencies in Europe and the United States. As part of the Development Centre's on-going work on governance in developing countries, this study hopes to contribute to improving its measurement and quality for the benefit of long-term development.

Louka T. Katseli

Director

OECD Development Centre

July 2006

ISBN 92-64-02685-1 © OECD 2006

Summary

A veritable explosion of interest in the quality of "governance" in the developing world is driving explosive growth in the use of governance indicators by international investors and both national and multilateral official OECD development co-operation agencies. Based on the maxim that you can only manage what you can measure, these decision makers seek to quantify the quality of governance in developing and emerging-market economies.

Among the hundreds of governance-indicator datasets that have emerged in response to this demand, the most widely used are composite perceptions-based indicators. More than users seem widely to perceive, however, even the most carefully constructed of these indicators lack transparency and comparability over time, suffer from selection bias, and are not well suited to help developing countries identify how effectively to improve the quality of local governance. Users — mainly people located outside developing countries — thus tend to use, and widely misuse, these indicators to compare the quality of governance both among countries and over time.

The perfect governance indicator will undoubtedly never exist. Still, the production and use of more *transparent* governance indicators will better serve the needs of both external users and developing countries seeking to improve the quality of local governance. Promising new developments in the "market" for governance indicators are on the horizon.

ISBN 92-64-02685-1 © OECD 2006

Introduction

The last 15 years have seen a veritable explosion of interest in the quality of "governance" in the developing world. Driving this growth are people who variously seek to monitor conditions in and/or assess prospects for diverse developing countries in terms of local political stability, investor-friendliness, economic growth or effective market size, poverty reduction, respect for human rights and long-term development. These people notably include international investors, national and multilateral providers of official development assistance, and development analysts and academics.

This growth of interest in the quality of governance has driven an equally explosive growth in the use of quantitative governance *indicators* in developing countries. A significant and rapidly growing number of international business and policy decisions directly rely on such indicators. A growing amount of analysis that influences broader perceptions, and often directly or indirectly shapes future decisions, does likewise. Yet numerous problems plague the use of governance indicators — some, unfortunately, considerably more serious than many users seem to realise.

This study seeks to clarify current trends in the use and misuse of governance indicators as those indicators are applied to developing countries. Chapter 1 highlights the main factors driving the growth of interest in governance. Chapter 2 examines some of the most widely used governance indicators. Chapter 3 takes a closer look at how key user groups actually use, and misuse, those indicators. Chapter 4 turns to a more in-depth analysis of what are undoubtedly the most carefully constructed and widely used governance indicators today — those produced by Daniel Kaufmann and his team at the World Bank Institute. Chapter 5 examines the two-way relationship between the quality of governance and economic growth to further illustrate the care with which even the best governance indicators must be used. Chapter 6 concludes with suggestions for future users and producers of governance indicators.

ISBN 92-64-02685-1 © OECD 2006

Chapter 1

Why all the Interest in Governance?

Summary

The quality of governance in developing and emerging-market economies has recently moved into the spotlight of international investors and official OECD development co-operation agencies, both national and multilateral, for a combination of reasons: *i)* the spectacular increase in international investment in developing countries; *ii)* the end of the Cold War; *iii)* failed development policy reforms in the 1980s and 1990s; and *iv)* a new awareness of the importance of politics in economic development and policy reform.

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 — in recent years.

International Investment

One is the spectacular growth of international investment in developing countries over the last 15 years. Foreign *direct* investment going to those countries, whether to create or acquire production capacities to serve local markets, or to serve global markets or the investors' home markets, has grown from an average annual net inflow of about $10 billion in the early 1980s, to over $67 billion in 1992-94 and over $150 billion since 1997. Equally spectacular and important for driving up interest in the quality of local governance has

ISBN 92-64-02685-1 © OECD 2006

been the growth of international *portfolio* investment in developing and "emerging market"[1] economies — notably by major pension funds and other large institutional investors — from net annual flows of below $2 million in the late 1980s to about $20 billion in bonds and another $26 billion in portfolio equity in the 1990s[2].

International investors' major newfound interest in the quality of governance in developing countries is thus in part simply a reflection of the spectacular increase in the value of their assets exposed to risk in those countries. Also very important, however, has been the sea change since the 1980s in economic policy orientation in the developing world, in favour of less interventionist, more market-oriented and more investor-friendly policy regimes. Competition among developing countries to attract foreign investment has also intensified greatly, giving added impetus to this sea change in policy orientation, and to perceptions among international investors of a significant degree of convergence or homogenisation of *de jure* policy regimes among developing countries seeking to attract investment. For many direct and portfolio investors alike, differences among developing countries' perceived credibility in policy *implementation*, and above all in the quality of their systems of *governance* (both political and corporate governance), which also weigh heavily in investors' perceptions of countries' policy credibility, have thus emerged as the single most important determinant of their investment-location decisions (Oman, 2000)[3].

End of the Cold War

A second set of phenomena driving the explosive growth of attention to 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 coloured by the latter's position in the bi-polar world created by the Cold War. US President Franklin Roosevelt's often-quoted remark about Nicaragua's ruthless dictator Anastasio Somoza — that "He's a bastard, but he's our bastard", because of the non-communist stronghold Somoza maintained in Central America — is emblematic of the attitudes and behaviour towards governments throughout the developing world until the end of the 1980s[4]. OECD governments, their national aid agencies and multilateral development organisations (including the OECD

ISBN 92-64-02685-1 © OECD 2006

and its Development Centre) sought to promote economic and social development in the "Third World" both to fight poverty and raise living standards, and to limit any temptation for developing countries to turn to communism. The focus was on trying to help governments to improve their policies *without* significantly questioning the quality of local governance *per se.*

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 recognise 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, to giving 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, and by the latter year 25 per cent of its lending was committed to law and public administration in borrowing countries[5].

Failed Policy Reform

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. Those policy reforms — reflected in the sea change in economic policy orientation noted earlier and sometimes referred to, at least in the Latin American context, as the "Washington consensus" — were spurred by a combination of factors. These notably included the onset of the Third World debt crisis in 1982, followed by the drying up of voluntary international bank lending to developing countries (especially the "sovereign" lending that had grown spectacularly to recycle petro-dollars in the wake of the 1973 oil shock). They also included the sustained decline in commodity prices and, in many countries, a collapse of local development banks together with that of import-substituting industrialisation strategies. The combined result was the markedly slowed growth that plagued much of the developing world from the 1980s and gave impetus to widespread policy reforms there during the 1980s and 1990s (Oman and Wignaraja, 1991).

ISBN 92-64-02685-1 © OECD 2006

Yet the resulting widespread policy shift during the 1980s and 1990s in favour of greater "market friendliness" in developing countries, notably in Latin America, Africa and South Asia — and actively encouraged by the multilateral financial organisations that found themselves in a strong position to do so through lending conditionality — ultimately proved relatively disappointing (Easterly, 2002). The ensuing debate over whether the relative failures are better explained by too much or rather by too little effective implementation of the recommended reforms[6] is less important for our purposes than is the general recognition, today, that the reforms were relatively unsuccessful. This recognition has 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 may go far to explain the relative reform failures of the 1980s and 1990s[7].

New Institutional Economics

Another key contributor to this understanding, especially but not only among mainstream economists, and 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. 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, 2005).

ISBN 92-64-02685-1 © OECD 2006

Notes

1. The term "emerging market economy" was reportedly coined in 1981 by Antoine W. van Agtmael of the World Bank Group's International Finance Corporation. International investors, especially banks and portfolio investors, now widely refer to the low- and middle-income countries where they lend and invest as such.
2. World Development Indicators Online, 2005, World Bank.
3. The OECD's Business and Advisory Committee thus noted in its November 2002, statement *Investment — BIAC Position on Incentives:* "The most important factor in creating favourable conditions to attract foreign direct investment is good governance. (…). If such conditions prevail, no special incentives are needed to attract foreign, or indeed domestic, direct investment."
4. The remark was made, of course, prior to the onset of the Cold War.
5. World Bank Annual Report 2004.
6. See, for example, Ortiz (2003), Lora and Panizza (2002).
7. 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).

ISBN 92-64-02685-1 © OECD 2006

Chapter 2

Sources of Governance Indicators

Summary

Users of governance indicators easily get lost in the jungle of hundreds of existing indicators. This chapter explains the most widely used governance indicators, which are composite perceptions-based indicators, and where to find additional information on the supply of 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 of the quality of local governance. The supply of governance indicators has grown significantly in response. Yet much of the new supply uses indicators whose origins precede the recent explosion of interest in governance. It is useful to look briefly at five of the most widely used such indicators, as illustrations, before we move in the next chapter to look more closely at how different users tend to use, and misuse, these indicators.

International Country Risk Guide

One of the most important governance indicators since its inception in 1980, certainly for international investors, is the privately-owned International Country Risk Guide (ICRG) rating system. Created in the wake of the costly

ISBN 92-64-02685-1 © OECD 2006

financial shock to international lenders caused by the fall of the Shah of Iran in 1979 (after the huge 1970s build-up of "sovereign" bank lending to developing countries, noted earlier), the ICRG is "designed to assess *financial*, *economic* and *political* risks in countries, and to compare them between countries [in order] to meet the needs of clients for an … analysis of the potential risks to international business operations"[1]. Country ratings are also designed to be comparable over time.

ICRG's *financial*- and *economic*-risk assessments rely entirely on *objective* measurements — however imperfectly they may be measured. These include the ratios of a country's foreign debt to its GDP, its foreign debt-service and its current-account balance to its exports, its net international liquidity to imports, its budget balance to GDP, and its levels of growth, inflation and GDP per capita.

ICRG's *political*-risk assessments, in contrast, rely entirely on its experts' *subjective* interpretations of pre-specified risk "components" whose pre-determined weights are made the same for all countries to facilitate comparison across countries and over time. The political-risk components comprise the following:

- a government's apparent ability to stay in office and to carry out its declared programme(s);
- socio-economic conditions that can fuel unrest and/or impinge on a government's actions (unemployment, consumer confidence, poverty);
- other factors affecting investment risks (contract viability, expropriation, constraints on profit repatriation, payments delays);
- internal and external political violence and conflict;
- corruption;
- military in politics;
- religious and ethnic tensions;
- democratic accountability;
- bureaucratic quality;
- strength and impartiality of the legal system and popular observance of the law.

ISBN 92-64-02685-1 © OECD 2006

While ICRG's composite indicator gives equal weight to the subjective perceptions of political-risk components on the one hand and to the objective financial- and economic-risk indicators on the other, the company also advises clients on means of adapting both the data and the weights "in order to focus ratings according to an investor's particular characteristics and needs". It provides ratings for 140 countries on a monthly basis and offers current, one-year and five-year assessments with projections framed in "best case" and "worse case" scenarios.

While complete monthly ratings with their underlying data are available to clients, academics can pay a significantly smaller amount for access to a "researchers' dataset" that comprises countries' annual averages on all the components of the political-risk assessment from 1984, excluding the most recent year.

Like all governance indicators, ICRG ratings are subject to non-negligible measurement errors. ICRG does not provide estimates of the size of those errors.

Freedom House

Another very important source of governance indicators is Freedom House, whose annual ratings of political rights and civil liberties in 192 countries are widely used by journalists, analysts and academics.

Freedom House is a private non-profit advocacy organisation founded in the United States in 1941 by prominent figures from both major US political parties to serve as a "steadfast opponent of dictatorships of the far left and the far right" and a "clear voice for democracy and freedom around the world"[2]. It is funded by a combination of US government support and tax-deductible grants and donations from private sources, which currently include over a dozen major foundations. It is governed by a Board of Trustees comprising some 36 prominent US politicians, former government officials, business and labour leaders, writers, academics and journalists, all "united in the view that American leadership in international affairs is essential to the cause of human rights and freedom".

In its core publication, "Freedom in the World", Freedom House rates both a country's *political rights* and its *civil liberties* on a scale of 1 to 7 ("1" is the highest (best) level and "7" the lowest), and the average of the two ratings is

ISBN 92-64-02685-1 © OECD 2006

used to designate the country's status as "free" (a score below 3), "partly free" (3 to 5) or "not free" (above 5). The ratings are calculated on the basis of in-house experts' subjective perceptions organised according to a checklist of questions reportedly inspired by the 1948 United Nations' Universal Declaration of Human Rights.

The checklist on *political rights* comprises ten questions divided into three categories: the electoral process; political pluralism and participation; and the functioning of government.

— The three questions on the *electoral process* ask whether or not the head of the executive and members of the legislative branches of government are "elected through free and fair elections", and whether there are fair electoral laws, equal campaigning opportunities, fair polling and honest tabulation of ballots.

— The four questions on *political pluralism* and participation ask if people have the right to organise in different political parties of their choice, if there is a realistic possibility for opposition parties to gain significant shares of the vote and take power through elections, if people's political choices are free from domination by the military, foreign powers, totalitarian parties, religious hierarchies or economic oligarchies, and if cultural, ethnic, religious and other minority groups have reasonable self-determination and participation in the political decision-making process.

— The three questions on the *functioning of government* ask whether freely elected representatives determine the government's policies, whether the government is free from pervasive corruption, and whether the government is accountable to the electorate between elections and operates with openness and transparency.

The checklist on *civil liberties* comprises 15 questions in four categories: the freedom of expression and belief, people's rights to associate and organise, the rule of law, and personal autonomy and individual rights.

— The four questions on *freedom of expression* and belief ask if there are free and independent or pluralistic media, if public and private expression of religion is free and there are free religious institutions, if there is academic freedom and the educational system is free of extensive political indoctrination, and if there is open and free private discussion.

ISBN 92-64-02685-1 © OECD 2006

— The three questions concerning *associational and organisational rights* focus on the freedom of assembly and demonstration, on the freedom of political organisation (including both political parties and civic or *ad hoc* organisations), and on whether there are free trade unions, peasant organisations, professional and other private organisations, and effective collective bargaining.

— The four questions on the *rule of law* ask if there is an independent judiciary, if the rule of law prevails in civil and criminal matters and the police are under direct civilian control, if people are protected from police terror, unjustified imprisonment, exile or torture (whether by groups that support or oppose the system) and from war and insurgencies, and if the population is treated equally under the law.

— The four questions on *individual rights* ask if there is freedom from indoctrination and excessive dependency on the state or there are state controls on travel, choice of residence or employment, if citizens have the right to own property and establish private businesses without undo influence by government officials, the security forces or organised crime, if there are personal social freedoms including gender equality, choice of marriage partners and size of family, and if there is equality of opportunity and the absence of economic exploitation.

For each of the 192 countries (plus a number of disputed territories) it currently rates, Freedom House publishes annually both ratings and the country's status as "free", "partly free" or "not free". It does not however make available to the public a country's scores on specific questions or groups of questions on the checklists. Nor are the data or ratings fully comparable over time, due both to periodic changes in the methodology and, presumably, to changes in the group of experts whose perceptions determine the ratings.

Nor does Freedom House provide estimates of the size of measurement errors embodied in its ratings (although, as for all governance indicators, these errors are certainly non-negligible).

Transparency International

Probably no governance indicator attracts more media attention than the Corruption Perceptions Index (CPI) published annually since 1995 by Transparency International. It is also widely used by investors, donors, analysts and academics.

ISBN 92-64-02685-1 © OECD 2006

The subject of corruption was practically taboo during the Cold War. Development agencies hardly discussed it, multilateral financial organisations largely felt they had to close their eyes to it, and the private sector widely saw it as an unpleasant and often costly but unavoidable part of trying to get things done in many parts of the world. This was still the case when Transparency International (TI) started as a small NGO in 1993. The catalyst in TI's creation was Peter Eigen, a former World Bank official with experience in Africa and Latin America who had argued unsuccessfully that the Bank should address the problem in its programmes. When Jim Wolfensohn became the Bank's President in 1995, convinced that corruption was an economic issue with a significant and direct negative impact on the effectiveness of the Bank's development programmes, he put corruption on the agenda and decided to work with TI to develop an anti-corruption strategy for the Bank. That same year TI developed the CPI to express the relative degree of corruption perceived in a country by the domestic and international business communities[3]. The CPI attracted massive global attention and helped to put corruption on the global development agenda[4].

The CPI can be understood as a survey of surveys. It is constructed by compiling the results of different surveys of perceptions of resident and non-resident business people and expert assessments in order to provide a snapshot of perceptions of the degree of corruption prevalent in a country, and then ranking the countries covered. The 2005 CPI ranked 159 countries based on the results of 16 surveys and expert assessments undertaken by 10 different organisations between 2003 and 2005. A country's CPI score (between "10" for the least corrupt and "0" for the most corrupt) is made public together with the number of surveys on which the score is based and an estimated "confidence range" of possible values of the CPI score depending on the estimated degree of measurement precision. Countries with fewer than three surveys or expert assessments are excluded — which means that many countries, including some among the most corrupt, are excluded for lack of perception data.

Year-to-year changes in a country's rank thus result not only from changes in perceptions of corruption in the country itself — whether because corruption has actually changed, or because subjective perceptions of it changed — but from changes in CPI's country sample base and methodology. Some sources are not updated and must be dropped, while new sources are added. Over time, with differing respondents and slightly differing methodologies used to

ISBN 92-64-02685-1 © OECD 2006

construct the CPI, a change in a country's score may thus be due to the fact that different viewpoints have been collected and different questions asked, rather than because of any change in the reality of corruption in the country.

While CPI scores are published annually, year-to-year comparisons of scores are thus hazardous. Nor are the disaggregated survey data — some of which are from commercial sources — made publicly available.

The World Bank

The World Bank produces two sets of governance indicators of major importance for our purposes. One, to which we return again in Chapter 4, is published bi-annually since 1996 by Daniel Kaufmann and his colleagues at the World Bank Institute. The other is the Country Policy and Institutions Assessments (CPIAs), which are produced annually by the Bank's own staff, i.e. its country teams, to assess the quality of Bank borrowing countries' policy and institutional frameworks for fostering poverty reduction, sustainable growth and effective use of development assistance. These Assessments have been used since 1977 to help guide the allocation of interest-free loans and grants by the Bank's IDA (International Development Association) to the poorest countries. In the past, a country's CPIA results were not made available to the public, however, and only recently have governments themselves, whose policies are assessed in a particular CPIA, come to be informed of the numerical ratings on a confidential basis.

The criteria used in the CPIAs have also evolved over the years, in response to new analytical insights and lessons the Bank feels it has learned from experience. Currently they comprise 16 criteria divided into four clusters:

— an *economic-management* cluster (comprising three specific criteria: macroeconomic management, fiscal policy, and debt policy);

— a cluster on *structural policies* (comprising three criteria: trade policies, financial-sector policies, and the business regulatory environment);

— a cluster on *policies for social inclusion and equity* (with five criteria: gender equality, equity of public resource use, building human resources, social protection and labour, and policies and institutions for environmental sustainability); and

ISBN 92-64-02685-1 © OECD 2006

— a *public-sector management and institutions* cluster (with five criteria: property rights and rules-based governance, quality of budgetary and financial management, efficiency of revenue mobilisation, quality of public administration, and transparency-accountability-corruption in the public sector).

The Bank's country team gives a score of 1 to 6 to a country for each of the 16 criteria, and gives each cluster the same weight (i.e. the criteria are not equally weighted) in producing the overall country assessment. The *public-sector management and institutions* cluster serves as a major input for the so-called "governance factor" which plays a critical role, in addition to the country's overall CPIA rating, in the allocation of Bank funds.

To enhance consistency of ratings across countries, the Bank now provides assessment teams with detailed questions and definitions for each of the six rating-levels; a Bank-wide process of rating and vetting a dozen "benchmark" countries is undertaken first. A Bank-wide review of all country ratings is also carried out before they are finalised.

Governments, as noted earlier, have recently been informed of the assessment process, which is increasingly integrated into processes of Bank-government dialogue. Starting in the summer of 2006, with the 2005 CPIA ratings, the Bank discloses to the public the numerical rating for each criterion, whereas previously assessment outcomes were "disclosed" only by grouping countries into quintiles according to the level of their results[5].

World Bank Institute

The most comprehensive publicly available set of governance indicators is published by the World Bank Institute. Available since 1996, these indicators are also the most widely quoted and widely used governance indicators in the media, academia and among international organisations. Along with Transparency International's CPI, they have played a leading role in putting governance on the agenda in developing countries[6].

Produced by the WBI's Daniel Kaufmann, originally with co-authors Aart Kraay and Pablo Zoido-Lobatón (hence widely referred to as the "KKZ" indicators) and now also Massimo Mastruzzi, this set of indicators was created in response to four inter-related concerns[7]. One was the apparent lack of robustness of cross-country comparisons using different individual data

ISBN 92-64-02685-1 © OECD 2006

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.

The WBI indicators are composite indicators of each of six aspects of governance: *i) Voice and Accountability*; *ii) Political Stability*; *iii) Government Effectiveness*; *iv) Regulatory Quality*; *v) Rule of Law*; and *vi) Control of Corruption*. The six indicators are composite in the sense that they are constructed from hundreds of existing perception indicators derived from 37 different data sources produced by 31 different organisations — including the ICRG, Freedom House, the World Bank (CPIAs) and most of the sources used by Transparency International for its CPI — as shown in Box 4.1 and Figure 4.1 in Chapter 4.

The country coverage of the KKZ indicators is very large — between 204 and 207 in 2004, depending on the indicator — thanks to the large number of sources used. Unavoidable measurement errors mean, however, that the indicators often cannot be used reliably to differentiate between levels of governance quality across countries. The authors provide statistical confidence intervals for each country's score on each indicator in a given year (see Figure 4.1 in Chapter 4), and only in the case of countries whose scores differ by so much that their confidence intervals do not overlap can one consider the difference between them to be meaningful (i.e. statistically significant under the assumptions made[8]).

The methodology used and the changing composition of the indicators over time further mean that the indicators cannot be used reliably to compare levels of governance over time, be it in a given country or among countries. When comparing scores over time, only if the change in or difference between scores is large enough that the scores' confidence intervals do not overlap can the change or difference be considered meaningful.

The measurement errors reflected in the scores' confidence intervals are unavoidable in the construction of governance indicators, as noted earlier. Kaufmann and his team, together with Transparency International, are, however, the only major producers of governance indicators who clearly highlight in all their publications the importance for users of taking these measurement errors into account. In doing this they provide an important

ISBN 92-64-02685-1 © OECD 2006

service: every attempt to quantify and compare governance levels inevitably involves measurement errors whose significance for users should be much more widely acknowledged and explained by producers.

Nor, once again, are the disaggregated data used to produce the KKZ indicators all publicly available[9].

Finding Your Way through the Jungle of Governance Indicators

Beyond these five sets of governance indicators, which are most widely used today by international investors, donor agencies and development analysts, there are of course many others. Indeed, by one recent estimate, there are now some 140 user-accessible sets of governance indicators, comprising literally thousands of individual indicators (World Bank Institute, 2006). Their proliferation has led in turn to the production of several governance-indicator "guides" and "inventories" that provide valuable "how to use" and "where to find" information on many of these indicators. We list six of these guides in Box 2.1.

These guides usefully distinguish between governance indicators that are perceptions-based (including those presented above) and indicators that are constructed from objective facts. The latter include indicators based, for example, on data on the existence or non-existence of specific anti-corruption laws, or of a corruption-prosecution agency, data on the number of legal prosecutions for corrupt acts, data on the existence or non-existence of regulations that make it more difficult to fire employees, or on the average cost of doing so, data on the number of procedures required legally to start a new business, or the average cost and time required to do so, data on the time it takes to acquire a new telephone line, data on voter turnout, etc.

The distinction between perceptions-based and facts-based indicators is important, not least because facts-based indicators are replicable and in this sense are more transparent for users than are perceptions-based indicators. Yet it would be a mistake to believe that facts-based governance indicators are necessarily more objective than perceptions-based indicators. Both the choice of facts used and, above all, the *interpretation* of how variations in those facts tend to affect the quality of governance mean that facts-based governance indicators embody a significant degree of subjective judgement in their construction — as do perceptions-based indicators, of course, in the very data

ISBN 92-64-02685-1 © OECD 2006

they use as inputs. Rather than seeing facts-based indicators as inherently more objective than perception-based indicators, in other words, users should understand perceptions-based and facts-based indicators as potentially useful *complementary* sources of information[10].

It nevertheless remains the case that international investors, donors and decision makers as a whole tend today to rely primarily on perceptions-based governance indicators. Two reasons seem largely to explain this tendency. One is that the data required to construct facts-based indicators are often lacking for developing countries, or the numbers that exist for those countries are perceived as lacking credibility. The other is that the data used to construct facts-based indicators often reflect only formal *de jure* realities, but these do not reflect *de facto* realities which are often informal and unwritten but nevertheless determine, much more than formal *de jure* realities, the true quality of governance in a country.

Thus, for example, the existence of specific anti-corruption laws does not necessarily imply lower *de facto* corruption in one country compared to another that does not have those laws, just as the formal creation of a corruption-prosecution agency may or may not reflect the seriousness with which a country actually prosecutes corruption. Similarly, a much larger number of legal prosecutions for acts of corruption in one country compared to another may just as easily reflect a higher or a lower level of corruption in the first country compared to the second. And, of course, to attribute better governance scores to countries whose regulations make it easier to fire workers — as does the World Bank's "Doing Business" set of indicators, for example — implies a significant degree of subjective judgement on the part of those who construct this facts-based indicator.

ISBN 92-64-02685-1 © OECD 2006

Box 2.1. **Governance Indicator Guides and Inventories**

OECD's Metagora: Metagora has developed a prototype of an online inventory of initiatives by local, national and regional organisations to measure human rights, democracy and governance. Placing particular emphasis on developing countries, the inventory is designed to provide a full description (topics, methods, budget, etc.) of each initiative it captures, along with information on the institutions and experts involved in their implementation, and links for accessing related publications and available technical documents. The inventory is designed as an open-ended tool; any person or institution initiating a relevant measuring project will be able to fill in an electronic questionnaire that will subsequently be controlled and registered into the database. http://www.metagora.org/html/activities/act_inventory.html

UNDP's "Governance Indicators": A User's Guide: Produced by UNDP's Oslo Governance Centre in collaboration with the European Commission, this guide provides direction for the non-specialist user on where to find and how to use free-of-charge sources of governance indicators. http://www.undp.org/oslocentre/docs04/UserGuide.pdf

World Bank Institute: This downloadable WBI inventory provides basic information, including the web link (or email address of the developer if no web link was found), for 140 sets of governance indicators, both commercial and free of charge. http://www.worldbank.org/wbi/governance/govdatasets/

The Human Rights Centre at the University of Essex: Published in 2003 under the title "Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance", the final report of this project aims to *i)* identify and analyse the main initiatives to develop indicators for measuring democracy, human rights and good governance by academics, inter-governmental organisations and non-governmental organisations; *ii)* evaluate the strengths and weaknesses of those initiatives; and *iii)* give recommendations on priority setting and basic orientations for developing related governance indicators. It was commissioned by the Statistical Office of the Commission of the European Communities (Eurostat). http://www.oecd.org/dataoecd/0/28/20755719.pdf

World Peace Foundation: Marie Besançon's Report "Good Governance Rankings: The Art of measurement" (2003) describes and analyses sources of governance indicators. It draws on the results of an expert meeting held at the John F. Kennedy School of Government, Harvard University. http://bcsia.ksg.harvard.edu/BCSIA_content/documents/WPF36Governance.pdf

Munck, G. and Verkuilen, J. (2002), also listed in our bibliography, provides a valuable and widely cited review and critique of democracy data. Their guidelines for aggregation and measurement are applicable to all governance data sets.

ISBN 92-64-02685-1 © OECD 2006

Notes

1. See http://www.icrgonline.com (our emphasis).
2. http://www.freedomhouse.org.
3. See http://www.transparency.org.
4. Also contributing to putting corruption on the global agenda during this period were the 1997 OECD Anti-Bribery Convention, and the major international conference on "Fighting Corruption in Developing Countries and Emerging Economies: The Role of the Private Sector" organised in Washington D.C. in 1999 by the OECD Development Centre with the support of the US Agency for International Development, the Center for International Private Enterprise, the MacArthur Foundation and PriceWaterhouseCoopers.
5. See Gelb, Ngo and Ye (2004) for estimates of the measurement error inherent in the CPIA.
6. See press coverage both within and outside developing countries at: http://www.worldbank.org/wbi/governance/press-2004indicators.html.
7. We thank Aart Kraay for this information.
8. See Chapter 4.
9. We were pleased to learn in June 2006, after benefiting from comments and criticism from Aart Kraay and Daniel Kaufmann on an earlier draft of this study, that the World Bank Institute has decided to begin disclosing countries' scores on each of the 37 data sources from which the composite KKZ indicators are constructed. We consider this an important improvement. It is perhaps worth noting, however, that most of these sources are themselves composite indicators, constructed from experts' and households' answers to survey questions or checklists, and that the user of KKZ indicators will not have access to countries' scores on all of these (e.g., Freedom House indicators are used as sources for the KKZ indicators, and Freedom House does not make available to the public a country's scores on specific questions or groups of questions on the checklists, as noted previously).
10. See Knack, Kugler and Manning (2003) for a discussion of policy-relevant facts-based indicators.

ISBN 92-64-02685-1 © OECD 2006

Chapter 3

Uses of Governance Indicators

Summary

User groups — mainly international investors, official national and multilateral aid agencies, and development analysts and academics (i.e. mostly people located outside developing countries) — tend to use — and widely misuse — governance indicators to compare the quality of governance both among countries and over time in their decision-making processes.

The primary direct users of governance indicators, besides journalists, are international investors, aid donors and academics. Each group tends to use — and to misuse — governance indicators in specific ways.

International Investors

Private capital flows to developing countries comprise three principle types: foreign *direct* investment (in which the non-resident investor has partial or total direct control over the management of the enterprise in which the investment is made); international *portfolio* investment (cross-border purchases of stocks, bonds and other securities where the investor has no such direct voice in the management of the invested enterprise); and international commercial *bank loans*. Together these flows amounted to an estimated $317 billion in 2004 (our most recent data), of which direct investment was about $132 billion, portfolio investment about $35 billion, and net flows from private creditors about $149 billion[1].

ISBN 92-64-02685-1 © OECD 2006

Because FDI usually constitutes a relatively long-term commitment of resources by the investor — funds invested in real assets are often not very liquid and thus relatively "hostage" to the success of the invested enterprise — foreign direct investors tend to spend much time and effort to compare countries they are seriously considering as potential investment locations. Political instability, weak rule of law, contempt for property rights, or a poorly functioning judiciary can easily discourage investors that perceive the risk of loss in a country as too high, or too difficult to gauge. The research departments of multinational corporate investors now widely construct or use governance indicators to try to assess the general country risk and governance situation in potential investment locations. A leading multinational corporation in the concession of water and sanitation infrastructure in developing countries interviewed for this study, for example, which describes the typical duration of its investments in developing countries as 20 to 25 years, gives particular attention to indicators of *political stability* and *the rule of law*.

Portfolio and other investors — which, together with speculators, we can call 'financial actors' because they tend to operate more exclusively within financial markets — also seek to estimate potential risks and returns, and are concerned about uncertainty, in the markets where they operate. As Keynes and many others since him have observed, financial markets' behaviour often depends more directly on market participants' *expectations* of what other participants will do ("herd behaviour") than on more objective economic "fundamentals". Some financial actors who believe a stock is overvalued, for example, will nevertheless be tempted not to sell it but to hold or even buy more of it in hopes of selling it at a still higher price, if they expect others are willing to do so. The level of risk associated with a stock price rises as the price moves away from the level that would be justified by the fundamentals. Precisely because of the relative importance of herd behaviour, moreover, investors' confidence usually changes not gradually or smoothly, but suddenly, and it is extremely difficult to predict when this turning point will occur. It is therefore crucial for financial actors to be able to assess markets' over-optimism or over-pessimism in order both to temper enthusiasm and to identify opportunities. France's *Caisse de Dépots et Consignations*, interviewed for this study, for example, thus reported that it undertook research to compare their in-house country-risk assessments based on fundamentals to the market's behaviour.

The 1980s "Third World debt crisis" drew attention to the fact that many internationally active banks failed to back up high-risk loans to developing countries with sufficient capital reserves to protect themselves in case of loan default or "non performance". This observation led to the "Basel I" agreement

ISBN 92-64-02685-1 © OECD 2006

in 1988, under the auspices of the Basel Committee for International Banking Supervision (and now part of national legislation in most countries), which stipulates that banks must hold an amount of capital on reserve, relative to the size of a loan, that varies according to the level of so-called *risk-weights* which the agreement attributes to different categories of borrower. Basel I primarily distinguished between OECD and non-OECD countries to determine risk-weights: Borrowing governments and central banks of OECD countries were assigned a *zero* risk-weight, and private banks a *20 per cent* risk-weight. While private banks in *non*-OECD countries, as in OECD countries, could also be assigned a 20 per cent risk-weight for short-term loans (i.e. those with a maturity of less than a year), governments and central banks in these countries were attributed a *100 per cent* risk-weight.

Of course, the higher is the risk weight assigned to the borrower, the larger is the amount of capital the lending bank must hold in reserve[2], and therefore the higher are the funding costs for the lender, which translate into higher interest-rates for the borrower. Basel I thus meant that for banks, loans to non-OECD borrowers, and especially long-term loans to non-OECD borrowers, cost considerably more than loans to OECD countries.

Dissatisfaction with Basel I, largely due to the arbitrary dichotomy between OECD and non-OECD countries and the failure to distinguish among borrowers of different risk levels among the latter, led to renewed discussions. Basel II, agreed in 2004, supersedes the simple dichotomy between OECD and non-OECD countries by allowing banks and other investors to use their own internally produced country-risk ratings to determine risk weights. Many now have their own country-risk-analysis experts, or departments, and — particularly important for our purposes — many of these in turn increasingly use governance indicators as a key element in their country-risk assessments.

This latter trend is new. Until recently, banks and international investors (including MNCs and other major direct and portfolio investors) that paid attention to country risk — as increasingly was the case — tended to rely on the "sovereign risk" assessments of the ability and willingness of sovereigns and companies to honour their financial obligations that are produced by the leading private rating agencies (notably Moody's, Standard and Poor's, and Fitch's). These ratings, whose components are not fully disclosed, are understood to rely primarily on such objective information as a country's GDP level and growth rate and the size of its fiscal and international accounts' balances. The higher the rating agencies' perception of a borrower's risk of default, the higher is the risk premium the borrower has to pay in the form of

ISBN 92-64-02685-1 © OECD 2006

higher interest rates, and the more likely are potential investors (direct and portfolio as well as creditors) to decide not to invest, or perhaps withdraw in the event of a downgrade.

Unfortunately, however, most country risk ratings failed to predict major financial crises over the last decade. Moreover, as Reisen (2003) explains, not only did they fail to predict the crises, they tended to lag behind the markets and, in doing so, to exacerbate the boom-bust cycle. The reason for this effect is that some of the key rating determinants, such as GDP growth and fiscal balances, are influenced by capital inflows and therefore not independent of investors' behaviour[3]. One of the striking features of the Asian crisis was thus the so-called "ratings crisis" (Jüttner and McCarthy, 2000), in which ratings downgrades — after the crisis had broken — seriously amplified the costs of the crisis, not only in individual borrowers but via contagion effects in other "emerging" economies as well.

The crises have thus been very costly for the borrowing countries, and their populations, as well as for their creditors. The output loss suffered by the crisis economies alone (Argentina, Brazil, Indonesia, Korea, Malaysia, Thailand and Turkey), for example, was estimated at over a trillion dollars — equivalent to $150 billion per year between 1995 and 2002[4].

Dissatisfaction with the traditional ratings systems has greatly reinforced international investors' attention to the quality of governance, and their demand for governance indicators, in developing countries. As one study puts it, "Whereas country risk analysts focused on debt ratios and growth rate indicators…consensus is emerging to place governance at the heart of the development process" (Bouchet *et al.*, 2003). Along with major direct investors, internationally active banks and asset managers now increasingly factor governance indicators into their investment decisions, and country risk ceilings[5]. The less developed the economy of a country is, furthermore, the more importance creditors and investors tend to give to these indicators[6]. Basel II, in allowing and encouraging international creditors to develop and use their own internal ratings-based systems of country-risk analysis, will certainly strengthen this trend.

Interviews undertaken for this study of 10 major internationally active banks and companies confirmed both the strong recent growth in such investors' predilection for using governance indicators in their lending and investment decisions, and the much greater emphasis they place on using them for their lending and investment decisions in developing as opposed to OECD countries. These interviews also highlighted investors' strong tendency to use

ISBN 92-64-02685-1 © OECD 2006

composite governance indicators, such as those produced by the World Bank Institute (the KKZ indicators) or Transparency International, which reduce several indicators for a country into a single composite score. The advantage of such composite indicators, as a country analyst for a large multinational bank explained, is that "they summarise a variety of sources" which he can use for the governance component in his ranking "without having to look at the disaggregated components".

The significant degree to which these indicators rely on investors themselves for information suggests, however, that the Minsky Tranquillity Paradox is never far away. As Bouchet *et al.* (2003) explain it, the Minsky Paradox refers to the fact that "after a long enough period of relative tranquillity, entrepreneurs and banks tend to become complacent about economic prospects. Little by little, they start to take more risk, going for more debt, and hence making the system more vulnerable". The reality of this "paradox" further amplifies the importance for investors not to follow blindly the herd, but actually understand the information conveyed (and not conveyed) in the governance indicators on which they increasingly rely.

Donors of Aid

Providers of official development assistance (ODA), both national governments and multilateral organisations, paid little attention to the quality of governance in recipient countries during the Cold War period, as explained earlier. Following the watershed decision by World Bank President James Wolfensohn in 1996 to reverse course, and give high priority to addressing corruption and bad governance as major barriers to development, the Bank undertook research showing a strong positive correlation between the quality of governance and the effectiveness of ODA in a recipient country. The Bank's study by Burnside and Dollar (1997), "Aid, Policies and Growth", became a foundation for aid allocation according to governance criteria.

Recent studies confirm that most donors now pay considerable attention to the quality of political governance in recipient countries when making their aid-allocation decisions. Berthélemy and Tichit (2004) found this to be the case in their study of more than 20 donors, and Burnside and Dollar (2004) did so on the basis of data from a large cross-section of developing countries. The latter study sums up the situation precisely: "In the 1980s, the amount of aid a country received was not correlated with institutional quality" — as

ISBN 92-64-02685-1 © OECD 2006

measured in the study by the Freedom House and ICRG indicators — whereas "in the 1990s the picture changed: countries with better institutions received significantly more aid. One standard deviation higher on the indices of rule of law and of democracy corresponded to 28 per cent more overall aid and 50 per cent more finance from the World Bank IDA facility…"[7].

As donors increasingly make the quality of governance in recipient countries an important criterion for aid-allocation decisions, they feel a growing need for governance indicators — not least to be able to base those decisions on consistent and *transparent* criteria. An informal recent survey of six official donors thus found broad support among them for the use of governance indicators in country-recipient selection. Donors' growing use of governance indicators also reveals, however, a number of serious problems or potential pitfalls associated with that use. Three sets of examples illustrate some of these problems and pitfalls.

1) World Bank CPIAs

The World Bank's Country Policy and Institutions Assessments (CPIAs), used notably by the Bank's concessional lending arm, IDA, is considered by many, together with the KKZ indicators, to be the most carefully constructed set of governance indicators. A major shortcoming for aid recipients, however, has been the CPIAs' lack of transparency. Reflected in the much-discussed difficulty for developing countries to challenge their CPIA scores, this lack of transparency limited a country's ability to target specific weaknesses that lay behind its score and thus effectively learn from the assessment process in order to improve the quality of local governance, and their CPIA score, over time. A further limitation of the CPIA is precisely the lack of comparability of scores over time.

Responding to widespread criticism of insufficient transparency, the World Bank started in 2006 to disclose to the public countries' CPIA scores.

2) Transparency International's CPI

Cited widely and frequently, in literally thousands of newspapers worldwide, Transparency International's Corruption Perceptions Index has played an invaluable role in originating and giving sustained impetus to what has become a global movement against corruption. The international shaming

ISBN 92-64-02685-1 © OECD 2006

that follows its publication of scores encourages a race to the top both among neighbours and globally, among countries at all levels in the league table[8]. Yet the CPI is also widely misused.

While Transparency International clearly asserts that the CPI is a *ranking* and cannot be used as a measure of national performance in the fight against corruption (Galtung, 2005), it is often (mis)interpreted by newspapers, and sometimes (mis)used by donors, as precisely such a measure[9]. The reason why the CPI is not a measure of corruption — and cannot reliably be used as a basis for aid-allocation decisions — is because year to year changes both in its methodology and in the list of countries it covers mean that it does not discriminate reliably either *i)* among countries with scores close to each other; or *ii)* between conditions of corruption, even in the same country, over time. Nor should it be forgotten that more than 50 countries, including many that are undoubtedly plagued by severe corruption, are not ranked on the CPI.

Put differently, the standardisation technique TI uses to produce the CPI emphasises the rank ordering of countries over internal reforms (or the lack thereof) in any country. This means that the CPI cannot be used as an indicator of progress to reward reform efforts in any country (and that a country's score will change even though corruption there remains unchanged). Changes in the composition of the sources and methodologies used, from year to year, mean that country-rank scores are not comparable over time. Thus, in its press releases, "TI warns against misinterpreting such arbitrary changes in the rank order of countries. Despite these warnings, media headlines frequently refer to changes in a country's rank order and the various caveats on TI's website remain largely unreported and widely misunderstood" (Galtung, 2005).

In its own words, TI "does not encourage the CPI to be used for decisions on aid allocation. Countries that are perceived as very corrupt…need help to emerge from the corruption-poverty spiral. If a country is believed to be corrupt, but is willing to reform, this should serve as a signal to donors that investment is needed in systemic approaches to fight corruption. And if donors intend to support major development projects in corrupt countries, they should pay particular attention to corruption 'red flags' and make sure appropriate control process are set up to limit graft" (Transparency International, 2004*a*).

3) Millennium Challenge Account's Use of KKZ Indicators

A third important illustration of potential pitfalls and problems associated with aid donors' use of governance indicators is the United States' use of such indicators in its newly created Millennium Challenge Account (MCA).

ISBN 92-64-02685-1 © OECD 2006

Announced in 2002 as a vehicle for up to $5 billion per year in new US official development assistance, the MCA relies heavily on composite governance indicators to determine the eligibility of recipients among low-income countries. Five of the 16 indicators it uses are KKZ indicators — *voice & accountability, government effectiveness, rule of law, control of corruption, regulatory quality* — with the one on *control of corruption* serving, moreover, as a decisive "in or out" criterion: countries that score below the median on this indicator are excluded from eligibility.

The care with which the KKZ indicators are constructed, and the related fact that they are so widely used, lead us to choose them for a more detailed analysis, in Chapter 4, of governance indicators' practical strengths and limitations. It is nevertheless important to note here that Kaufmann and his co-authors have criticised the MCA's announced use of one of their composite indicators as the basis for such an "in or out" eligibility rule in aid allocation. Their criticism is two-fold: *i)* the measurement errors embodied in their composite indicators are such that "for the majority of countries there is a non-trivial probability that they could be mistakenly classified in the bottom half of the sample"; and *ii)* fewer sources of the information used to construct the indicator are available for low-income countries, whose scores may be based on just one or two such sources, which increases the risk of misclassification of such countries (Kaufmann *et al.*, 2002, 2003). A third criticism, which we would add, is that whereas the MCA should reward improvements in actual governance practices, the KKZ indicators are statistically not designed to reliably compare levels of governance over time — as we explain in greater detail below in Chapter 4.

Repeated warnings by both the WBI team and Transparency International that their indicators should not be used to compare countries that have similar scores, or to make comparisons over time without due caution, nevertheless remain widely ignored. And their indicators are among the most carefully constructed and widely used. While donors contacted in the informal survey reported that governance indicators are never the only criteria they use in aid allocation, and that they are aware that problems exist with governance indicators, the importance of governance indicators for donors continues to grow and so, unfortunately, does their misuse. Paradoxically, while an important reason for this growth is that donors want to use governance indicators to increase the transparency of their budget-allocation processes, as well as their consistency, the un-transparent nature of the composite indicators on which they tend to rely has rather the opposite effect.

ISBN 92-64-02685-1 © OECD 2006

Analysts and Academics

The potential for development analysis — by academics and others — to influence decision makers is considerable. The Burnside and Dollar (1997) study, cited earlier, illustrates: It became the cornerstone for a 1998 World Bank report that US officials cite as "the key study" supporting the creation of the Millennium Challenge Account, for example, and the official British and Canadian aid agencies are reported as saying that solid research backs the view that aid reduces poverty in countries that are well governed (Eviatar, 2003). Correlation and regression analyses that use corruption indicators have similarly been crucial in putting the corruption issue on the international agenda. As Transparency International's former head of research explains, "the heuristic and political function of these studies has been considerable. These studies provided the prima facie case required, for example, by the Bretton Woods Institutions to legitimise their commitment to anti-corruption. This was a topic they had hitherto explicitly avoided as being 'political' and beyond their remit for 50 years of their existence" (Galtung, 2005).

Yet researchers' use of governance indicators is widely plagued with pitfalls. The nature and limitations of composite governance indicators, which their users often seem not fully to grasp, or admit, unfortunately weaken the rigour and thus the credibility of many studies. To illustrate, we turn again to examples from the use of CPIAs, the CPI and Freedom House indicators, and the KKZ indicators.

1) CPIAs

World Bank employees are allowed to use CPIAs for research provided they do not disclose individual country scores. Thus, for example, Kraay and Nehru (2004) use CPIAs from 1977 (extrapolated back to 1970) through 2001 and find a significant inverse correlation between the quality of a country's policies and institutions, on one hand, and its probability of debt distress, on the other. The Bank's own Operations Evaluation Department has nevertheless warned against interpreting internal Bank research as finding that "good policies" as measured by CPIAs explain growth performance (World Bank, 2001), since it doubted that "management succeeded to gauge policy measures taken, and not development outcomes, which are not fully within the control of governments in developing or any other countries. […] Perhaps the warning should be circulated (or re-circulated) to World Bank researchers, as they continue to use the CPIA as an explanatory variable in econometric exercises" (Herman, 2004).

ISBN 92-64-02685-1 © OECD 2006

2) CPI and FH Indicators

Similar comments focus on other indicators: "Transparency International's CPI ratings are not comparable from year to year and small shifts in the annual rankings are not meaningful. Nevertheless, CPI rankings are often misused as a causal variable for cross-national time series studies. Likewise, Freedom House's 'Freedom in the World' Index gives an overall perception of the state of freedom within countries, but is not necessarily useful as an identifier of causal mechanism for failures. Both of these ratings systems are inherently subjective and not grounded in explicit theory" (Besançon, 2003).

The missing theory is also reflected in very different interpretations of the same indicator. Different analysts use the same Freedom House ratings, for example, as a proxy for everything from "democracy" to "institutional framework" to "human rights" to "rule of law" to "governance" (Landman and Häusermann, 2003).

3) KKZ Indicators

Kaufmann and his co-authors explicitly caution against using their indicators for ranking purposes because of the measurement errors embodied in their indicators (errors that are not unique to their indicators, it must be emphasised). Nor do they aggregate their six composite indicators, or a subset thereof, to create an overall composite indicator — they always use them separately — because, in their own words, of "non-trivial issues when constructing one composite governance indicator for a country" (Kaufmann *et al.*, 2005*b*). Many important studies nevertheless (mis)use the KKZ indicators precisely in these ways. Of the countless examples we could provide, three suffice.

One is important econometric analysis of the effects of countries' governance infrastructure on FDI inflows and outflows by Globerman and Shapiro, as published in such influential journals as *World Development* (2002) and the *Journal of International Business Studies* (2003). These authors find that countries which fail to achieve a minimum threshold of effective governance are unlikely to receive much FDI, and that above that threshold the quality of a country's governance infrastructure is an important determinant of the amount received. The strength of these findings and corresponding possibility to draw policy inferences from them is weakened by the fact that the authors use an aggregate of the six KKZ indicators as a proxy for governance infrastructure.

The Washington D.C.-based Center for Global Development provides another important example with its "Index of Aid Effectiveness" (Roodman, 2005). Donor countries score higher on this Index by giving more aid to

ISBN 92-64-02685-1 © OECD 2006

countries with relatively good governance (as well as by giving aid to countries with high poverty, and they are penalised for giving tied aid, and for receiving debt repayments from poor countries). Here again, unfortunately, the authors use the aggregate of the KKZ indicators to judge the quality of governance in recipient countries that determines donors' scores on the Index, which weakens the value of including governance in the Index.

It is understandable that users would like to have an overall indicator of governance. However, taking a simple average of the six KKZ indicators and using it in studies such as those mentioned above is problematic for at least two reasons. First, taking a simple average means loosing all the statistical advantages of the aggregation method KKZ use to produce both the composite indicators and the corresponding confidence intervals (based on estimates of measurement error embodied in the indicators) for countries' scores on a given indicator, as we explain in Chapter 4. Second, it is no more appropriate to use an aggregate "governance" score that combines the different KKZ indicators into a single number for a given country than it is to aggregate the quality of apples and the quality of oranges: if the quality of apples is very bad and the quality of oranges is very good, saying that the quality of fruit is satisfactory would mask the respective quality differences in the different types of fruit.

To illustrate the second reason, consider a comparison of China and India, which have similar scores if one looks at an aggregate of each country's scores on all six KKZ indicators for 2004. Yet China scores in the upper half of all countries on "Government Effectiveness" and in the lower quarter of all countries on "Voice & Accountability", whereas India scores in the middle of all countries on both these indicators. The aggregate hides potentially important differences between the two countries in their scores on different KKZ indicators, yet provides no meaningful overall indicator of governance with which to compare the two countries.

Similarly, if a country scores very well on the majority of KKZ indicators but very badly on one of them, the country's overall aggregate still presents a favourable picture of the country's "governance". Israel and Oman illustrate: They have similar aggregate scores that lie in the upper half of all countries, but Israel's score on "Political Stability" and Oman's score on "Voice & Accountability" both lie in the lowest quarter of all countries. Users who rely on an aggregate of the six KKZ indicators would be easily induced to believe that the quality of governance (fruits) is fine, and comparable, in the two countries without realising that there is a serious problem with the "Voice & Accountability" apple in Oman and the Israeli "Political Stability" orange.

ISBN 92-64-02685-1 © OECD 2006

It bears repeating, with emphasis, that each of the six KKZ indicators is already extremely complex on its own. This complexity, which reflects the large number and diversity of existing indicators used to construct each KKZ indicator, as we explain in greater detail in the next chapter, means that the substantive content or true "meaning" of each composite indicator is already lacking in transparency and very difficult to decipher or interpret in real governance terms. This lack of transparency is only multiplied for an aggregate that combines several of the KKZ indicators — to the point that it becomes very difficult to attribute any substantive "governance" meaning to such an aggregate.

Our third example is the landmark UNDP Programme on Governance in the Arab Region, which has played an invaluable role in bringing attention to bear on the problems of governance in Arab countries. The problem — which does not diminish the importance of the Programme in other respects — is that it uses KKZ and CPI scores to compare the quality of governance among countries whose scores are much too close to each other, compared to the measurement errors inherent in the calculation of those scores, to allow for any such comparison to be meaningful. Not only does the Programme make such cross-country comparisons, it does so over time — and even provides a tool on its webpage that encourages users to do likewise. Unfortunately, neither such comparisons among countries with similar scores nor their comparison over time is analytically sound — as we explain in greater detail in the next chapter.

To sum up this chapter, then, it is clear that international investors, official donors, and development analysts and academics all show a strong propensity to use composite governance indicators to discriminate among countries and identify trends over time in ways the indicators are not designed to permit. Investors do so mainly in country-risk analysis, donors for aid-allocation purposes, and academics for regression analysis. The danger, indeed the likelihood, is that major business and policy decisions are being made on false premises.

Our focus on international investors, donors and academics also reveals another important fact: governance indicators are very largely used by *external* observers and decision makers, as opposed to domestic groups and policy makers in most developing countries. The latter groups are increasingly aware of the importance of the quality of local governance, not only for attracting foreign investment and aid flows but for strengthening their economies and long-term development processes as a whole. Yet the production of governance indicators that can be used by domestic groups to clarify the nature of the obstacles to better local governance, and to monitor their progress in overcoming them, is a challenge that remains largely unsolved. We return to this issue in Chapter 6.

ISBN 92-64-02685-1 © OECD 2006

Notes

1. By way of comparison, total official development assistance that year was $78 billion, or about one-fourth the size of private flows (whereas ODA was about *three times* the size of private capital flows the early 1960s). Data on investment flows are from the Institute of International Finance (2005), Capital Flows to Emerging Market Economies, 24 September 2005.

2. The capital adequacy ratio is 8 per cent. This means that if the risk weight is 100 per cent, for example, then 8 per cent of the amount of the loan must be held by the lender on reserve.

3. For example, when investment pours into a country it tends to stimulate growth of GDP and to improve fiscal balances, both of which serve to sustain or even improve the country's risk ratings even though the actual degree of country risk may be high; the good ratings swell investors' confidence and attract more capital inflows; etc. The spiral of optimism may continue until some warning signal causes a market collapse followed by a ratings downgrade, which in turn amplifies the collapse as investors take their money out and rating agencies react with further downgrades.

4. Griffith-Jones and Gottschalk (2006).

5. Many financial actors seek to manage their exposure to risk by imposing on themselves a "country-risk ceiling" that stipulates a maximum amount that can be invested or loaned to counterparts — thus exposed to risk — in a given country, no matter how good the apparent risk profile of a potential new borrower or investment project in the country. Because financial actors do not see most OECD countries as risky, they mainly apply these ceilings to developing countries (see, for example, UBS, 2004 -5).

6. According to our interviews with the Institute of International Finance (IIF), there are serious indications that banks, investment banks and asset managers factor governance indicators into their investment decisions and country limit settings, and that the lower the level of a country's economic development, the greater the importance such financial actors tend to give these indicators.

ISBN 92-64-02685-1 © OECD 2006

7. Similarly, Dollar and Levin (2004) find that "the same group of multilateral and bilateral aid agencies that are very policy focused are also very poverty focused. The donors that appear high up in both rankings are the World Bank's IDA facility, the IMF's Enhanced Structural Adjustment Facility, Denmark, the United Kingdom, Norway, Ireland and the Netherlands. Japan scores highly on the policy/institutional selectivity index but far down the poverty selectivity index (reflecting the focus of its aid on Asian countries that are relatively well-governed but not in many cases very poor)".

8. The *Financial Times* thus noted in 2004, "Governments as diverse as Papua New Guinea, Cameroon and Bosnia-Herzegovina have started or stepped up anti-corruption programmes as a result of publicity generated by the Index… South Korea has even pledged to reach position 10 or above by 2007 — a tall order, as Seoul was ranked 50th last year".

9. At least one major donor is known to have decided to stop funding a country on the basis of its CPI score, for example.

ISBN 92-64-02685-1 © OECD 2006

Chapter 4

In-depth Analysis of KKZ Composite Indicators

Summary

Probably the most carefully constructed and certainly among the most widely used governance indicators are those produced by Daniel Kaufmann, Aart Kraay and their team at the World Bank Institute. We therefore take an in-depth look at the strengths and weaknesses of these particular indicators. Problems associated with their construction or use, of which users seem widely unaware, include: *i)* the likelihood of correlation of errors among the 37 sources from which the composite WBI indicators are constructed, which significantly limits the statistical legitimacy of using them to compare countries' scores; *ii)* their lack of comparability over time; *iii)* sample bias; and *iv)* insufficient transparency.

The preceding overview of how governance indicators tend to be used, and often misused, points up a need for better understanding of the strengths and limitations of the indicators — perception indicators — that are currently used the most. We therefore take a closer look at what are probably the most carefully constructed and widely used indicators: the KKZ indicators. Our task is greatly facilitated by the very extensive documentation provided by the authors themselves both on the construction and on the limitations of their indicators.

ISBN 92-64-02685-1 © OECD 2006

Methodology

The KKZ indicators refer to a concept of governance that does not emerge from, or imply, a theory of governance. The authors define governance simply as "the traditions and institutions by which authority in a country is exercised" and then interpret governance as comprising three "dimensions", each of which is defined or "captured" by two specific composite KKZ indicators, as follows:

1) The process by which governments are selected, monitored, and replaced:

— *Voice and Accountability (VA)*: The extent to which citizens of a country are able to participate in the selection of governments. Includes indicators measuring various aspects of political process, civil liberties and political rights, the independence of the media.

— *Political Stability (PS)*: Perceptions of the likelihood that the government in power will be destabilised or overthrown by possible unconstitutional and/or violent means, including domestic violence and terrorism.

2) The capacity of the government to effectively formulate and implement sound policies:

— *Government Effectiveness (GE)*: Perceptions of the quality of public service provision, quality of bureaucracy, competence of civil servants, independence of the civil service from political pressures, credibility of the government's commitment to policies.

— *Regulatory Quality (RQ)*: The incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development.

3) The respect of citizens and the state for the institutions that govern economic and social interactions among them:

— *Rule of Law (RL)*: Success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions and the extent to which property rights are protected. Includes perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts.

ISBN 92-64-02685-1 © OECD 2006

— *Control of Corruption (CC)*: Perceptions of corruption defined as the exercise of public power for private gain. Including both petty and grand corruption and state capture.

Rather than starting from an explicit theory of governance, or of how key components of a system of governance interact to determine the quality of the system as a whole, in other words, each of the six composite KKZ indicators effectively determines, empirically, the meaning of "governance" embodied in the particular indicator. The six indicators are constructed from, and their meaning thus determined by, hundreds of existing indicators produced by others — all *perception* indicators — compiled from 37 different data sources, supplied by 31 different organisations, as described in Box 4.1. It is thus the perceptions embodied in hundreds of different indicators, produced for diverse purposes, which give content to the concept of governance that emerges from the KKZ indicators via the authors' choice of existing indicators and the techniques they use to aggregate those indicators, produced by others, into their own composite indicators.

To produce one of their six composite indicators — take "Rule of Law" as an example — KKZ start with existing perception indicators which they believe likely to contain information useful for assessing the quality of Rule of Law in different countries. They proceed in five stages (see Appendix I for the detailed procedure).

First, having identified existing indicators which they judge relevant for the construction of a given composite indicator, KKZ aggregate those existing indicators by source in order to calculate a single number for each source (i.e. they combine all existing indicators they use from a particular source to produce a single number for each country covered by the source). The aggregation method KKZ use for this stage is to calculate a *simple*, un-weighted, *average* of all the existing indicators they use from a source in order to produce the single number for the source[1].

ISBN 92-64-02685-1 © OECD 2006

Box 4.1. **Sources of Governance Data**
used to construct the KKZ indicators (37 sources)

Cross-Country Surveys of Firms (5 sources):

— *Global Competitiveness Survey*: Produced since 1996 by the World Economic Forum, a private non-profit organisation headquartered in Geneva, Switzerland, which brings together leaders from business, government, academia and the media, this survey compiles business executives' perceptions of countries in which they operate. Covers 104 developed and developing countries. Source for all 6 KKZ indicators.

— *World Competitiveness Yearbook:* Produced since 1987 by the Institute for Management Development, a non-profit research and educational foundation based in Lausanne, Switzerland, this survey analyses the competitive environment in 49 developed and developing countries based on both objective data and surveys of perceptions from over 4 000 local and foreign enterprises operating in the countries covered. Source for all 6 KKZ indicators.

— *Business Environment and Enterprise Performance Survey* (BEEPS): Produced jointly by the World Bank and the EBRD, this survey first gathered perceptions in 1999-2000 from over 4 000 firms in 22 transition countries on a wide range of issues concerning interactions between firms and the state. In its second round, conducted in 2002, it covered over 2 100 firms in 27 transition countries. Source for: GE, RQ, RL, CC.

— *World Business Environment Survey*: Managers of at least 100 firms per country in 80 developed and developing countries were the respondents to this survey on the business environment facing private enterprises conducted by the World Bank in collaboration with several other institutions during 1999-2000 (questions similar to the 1997 WDR survey which KKZ used to construct the 1998 version of their indicators). Source for: all 6 KKZ indicators.

— *Africa Competitiveness Report 1998*: Analysis by the World Economic Forum of the business climate in 23 African countries. Source for: PS, GE, RQ, RL, CC.

Cross-Country Surveys of Individuals (5 sources):

— *Voice of the People:* Initiated in 2002 by Gallup International, an association of mostly for-profit market research companies in almost 60 countries, registered in Zurich, Switzerland, this annual survey interviews citizens in 62 developed and developing countries with a view to understanding the opinion of today's world population on issues related to the environment, terrorism, global issues, governance and democracy. Source for: VA, PS, GE, RL, CC.

ISBN 92-64-02685-1 © OECD 2006

Box 4.1 (contd.)

— *Gallup International Millennium Survey*: 1999 Gallup International poll of 57 000 adults in 62 developed and developing countries on topics of an ethical, political and religious nature. Source for: VA, PS, GE, RL, CC.

— *Latinobarómetro:* Produced since 1996 by Latinobarómetro, a non-profit organisation based in Santiago, Chile, this public opinion survey covers 17 Latin American countries. Source for: VA, PS, GE, RL, CC.

— *Afrobarometer*: Launched in 1999, Afrobarometer is a non-profit joint enterprise of Michigan State University, the Institute for Democracy in South Africa and the Centre for Democracy and Development in Ghana. It reports the results of national sample surveys on the attitudes of citizens in 12 African countries towards democracy, markets and other aspects of development. Source for: VA, GE, RL, CC.

— *Latin America Public Opinion Project (LAPOP)*: Directed by Prof. Mitchell Seligson of Vanderbilt University with the financial assistance of the United States Agency for International Development, the LAPOP conducted "Democracy Surveys" in 2004 on citizen attitudes towards democracy, corruption and related subjects in 7 Central American countries and Mexico. Source for: VA, RL, CC.

Expert Assessments from Commercial Risk Rating Agencies (10 sources):

— *Business Risk Service:* Provided by Business Environment Risk Intelligence (BERI), a for-profit firm headquartered in Geneva, Switzerland, which supplies analysis and forecasts of the business environment in 50 developed and developing countries that it monitors three times per year, assessing 57 criteria separated into three indices. BERI convenes two permanent panels of about 150 experts from all over the world which provide country ratings on the basis of initial reports written by BERI analysts. Source for: PS, GE, RL, CC.

— *Quantitative Risk Measure in Foreign Lending:* BERI (see previous source) provides estimates of qualitative risk factors in credit exposure in 50 developed and developing countries out of 115 countries covered in its "Lender Risk Rating" service. Source for: PS, GE, RL CC.

— *Country Risk Review (CRR):* This quarterly publication provides expert assessments on 117 developed and developing countries. It is sold since 1996 by Global Insight's DRI, a for-profit US economic consulting and information company founded in 1973 which provides data, analysis, forecasts and expert advice to strategic planners, business and financial analysts and policy makers. Source for: PS, GE, RQ, RL, CC.

ISBN 92-64-02685-1 © OECD 2006

Box 4.1 (contd.)

— *International Country Risk Guide (ICRG)*: Produced since 1982 by the Political Risk Services (PRS) group, a for-profit affiliate of the Investment Business with Knowledge (IBK) company based in Syracuse, New York, which provides country information for international business. The ICRG provides assessments of political, economic and financial risks in 140 developed and developing countries based on the analysis of a worldwide network of experts. Source for: all 6 KKZ indicators.

— *EIU Country Risk Service and Country Forecasts:* Country risk ratings and two quarterly publications containing governance indicators for 120 developed and developing countries based on regular contributions from a global network of about 500 information-gatherers, checked by a panel of regional experts for accuracy, consistency and impartiality; the publications were launched in 1997 by the Economist Intelligence Unit (EIU), a for-profit organisation founded in 1949, based in London, which produces analysis and forecasts of the political, economic and business environment in more than 180 countries. Source for: all 6 KKZ indicators.

— *World Markets Online* (WMO)*:* An online subscription service updated daily which provides analysis of the conditions and risk for businesses in 202 developed and developing countries; produced by the World Markets Research Centre, based in London and established in 1996, which employs about 190 permanent staff. WMO has developed a risk rating system to compare and contrast countries' investment climates drawing on a worldwide network of information gatherers and analysts. Source for: all 6 KKZ indicators.

— *iJET security risk rating*: iJET monitors the world around-the-clock and provides data on 167 developed and developing countries. Based in Annapolis, Maryland, it is a for-profit travel consultancy founded in 1999 that alerts travellers, expatriates and decision makers to events and situations in real-time to help them avoid or minimise risk and travel disruptions abroad. Source for: PS.

— *Gray Area Dynamics*™: Provided by the Merchant International Group (MIG), a for-profit strategic research and corporate intelligence company established in 1982, headquartered in London, which offers services ranging from the identification to evaluation of risks, weaknesses and threats to corporations in non-domestic markets. Gray Area Dynamics*TM* assesses a range of events, activities and trends that impact upon business in 154 developed and developing countries. Source for: PS, GE, RQ, RL, CC.

ISBN 92-64-02685-1 © OECD 2006

Box 4.1 (contd.)

— *Political Economic Risk Consultancy* (PERC): Provides data from surveys of expatriate business managers' perceptions of corruption, the quality of the legal system, and the professionalism and reliability of the policy and judiciary in 14 developed and developing countries in East and Southeast Asia. Founded in 1976 and headquartered in Hong Kong, China, PERC is a for-profit company that specialises in strategic information and analysis for companies doing business in this region. Source for: CC.

— *Opacity Index*: This index of the lack of transparency in 35 developed and developing countries in 2000 provided by PricewaterhouseCoopers, a for-profit US-based professional services firm, was constructed using a team of economists, survey professionals, analysts and advisors. Source for: GE, RQ, CC.

Expert Assessments from NGOs, Think Tanks (12 sources):

— *Press Freedom Index:* Published since 2002 by the NGO Reporters without Borders for 138 countries, this index is based on surveys of journalists, researchers, legal experts and human rights activists worldwide. Headquartered in Paris, Reporters without Borders is an international organisation dedicated to the protection of reporters and respect of press freedom in the world. Source for: VA.

— *Index of Economic Freedom:* Launched in 1995, this annual index of economic freedom in 156 developed and developing countries is produced by the US-based Heritage Foundation in partnership with the Wall Street Journal. The Heritage Foundation is a non-profit research and educational institute, founded ion 1973, whose mission is to formulate and promote conservative public policies based on the principles of free enterprise, limited government, individual freedom, traditional American values, and a strong national defence. Source for: RQ, RL.

— *Freedom in the World:* Produced since 1955 (annually since 1978) by Freedom House, an American NGO created in 1941 to promote democratic values around the world, this publication estimates Political Rights and Civil Liberties in 193 developed and developing countries based on subjective expert assessments. Source for: VA.

— *Nations in Transit*: This Freedom House publication, launched in 1995, evaluates the progress in democratic and economic reform in 27 transition economies on the basis of country surveys written by Freedom House staff or consultants reviewed by academics and senior Freedom House staff. Source for: VA, GE, RL, CC.

ISBN 92-64-02685-1 © OECD 2006

Box 4.1 (contd.)

— *Countries at the Cross Roads*: This Freedom House publication, launched in 2004, covers 30 developing countries "at crossroads in determining their political future" whose performance it evaluates in terms of democratic governance. Source for: VA, RL, CC.

— *Cingranelli & Richards Human Rights Database:* Produced by the University of Binghamton, which draws from the US State Department's Country Reports on Human Rights Practices and Amnesty International's Annual Reports, this data set contains quantitative information annually since 1981 for 192 developed and developing countries on government respect for 13 internationally recognised human rights. Source for: VA, PS, RL.

— *Political Terror Scale*: Provides information on 192 developed and developing countries contained in Amnesty International's Annual Reports and the US State Department's Country Reports on Human Rights Practices coded by the University of North Carolina annually since the early 1980s. Source for: VA, PS, RL.

— *Bertelsmann Transformation Index (BTI):* In 2004, the Bertelsmann Foundation, established in 1977 and based in Germany, began publishing the BTI, a global ranking that analyzes and evaluates development and transformation processes in 116 developed and developing countries. Source for: VA, GE, RQ, RL.

— *Global E-Governance Index:* Brown University's Center for Public Policy compiles this Index based on their evaluation of official websites in 192 developed and developing countries for the presence of various features in those websites dealing with information availability, service delivery, and public access. Source for: GE.

— *Media Sustainability Index:* Introduced in 2002, with financial support from USAID, by the International Research & Exchanges Board (IREX), an international non-profit organisation specializing in education, independent media, internet development and civil society programmes, this Index looks at the entire media system in each of 18 developing countries in Southeast Europe and Eurasia. Source for: VA.

— *Index of Budget Transparency*: Based on panels of experts (legislators, media, academic experts, NGOs), this Index evaluates different aspects of governments' budgetary processes, such as citizens' access to budget information, citizen's participation and the credibility of institutions, in 10 Latin American countries. It is produced by Fundar, a Mexican NGO, together with leading NGOs in the countries covered. Source for: VA.

— *State Capacity Survey*: Covering 108 and 97 countries from assessments completed by 164 experts during 2000 and 2002, respectively (KKZ use those surveys for their 2000, 2002 and 2004 indicators), this Survey was developed in 1999 under the direction of Marc Levy of Columbia University. Source for: VA, PS, GE, RL, CC.

ISBN 92-64-02685-1 © OECD 2006

Box 4.1 (contd.)

Expert Assessments from Governments, Multilaterals (5 sources):

— *World Bank Country Policy and Institutional Assessments* (CPIAs): Produced annually since the late 1970s by World Bank staff assessing the quality of policy and institutional performance in 136 developing World Bank borrowing countries. While in earlier years assessments focused mainly on macroeconomic policies, they now include factors such as social inclusion, equity and governance. Source for: GE, RQ, RL, CC.

— *Transition Report, European Bank for Reconstruction and Development* (EBRD): The EBRD's annual *Transition Report* includes governance variables in 26 transition economies with subjective indicators based on a checklist of objective measures and the views of EBRD staff. Based in London, the EBRD is an international organisation that promotes private and entrepreneurial initiative in transition economies. Source for: all 6 KKZ indicators.

— *African Development Bank's Country Policy and Institutional Assessments*: Similarly to the World Bank's CPIAs, African Development Bank staff annually assess and produce indicators on the quality of their 50 African developing-country borrowers' policy and institutional performance in areas relevant to growth and poverty reduction. Source for: GE, RQ, RL, CC.

— *Asian Development Bank's Country Policy and Institutional Assessments*: Similarly to the World Bank's CPIAs, Asian Development Bank staff annually assess and produce indicators on the quality of their 26 Asian developing-country borrowers' policy and institutional performance in areas relevant to economic growth and poverty reduction. Source for: GE, RQ, RL, CC.

— *Progress towards Good Governance in Africa, UNECA:* The United Nations Economic Commissions for Africa (UNECA) produces Africa Governance Indicators for 23 African countries that are the result of a study to measure and monitor progress of governance in Africa, published in "Progress towards Good Governance in Africa", which incorporate assessments based on expert panels, population surveys and factual data gathering. Source for: VA, GE, RQ, RL, CC.

KEY: VA = Voice & Accountability GE = Government Effectiveness PS = Political Stability RQ = Regulatory Quality RL = Rule of Law CC = Control of Corruption

A more detailed description together with a list of the indicators used from each source for KKZ can be found in Kaufmann *et al.* (2005*a*).

ISBN 92-64-02685-1 © OECD 2006

Thus, in the case of "Rule of Law", KKZ use existing indicators from 24 sources in 2004. One of those sources is called "World Markets Online" and is produced by the organisation World Markets Research Centre (see Box 4.1). Among the several indicators produced by this organisation, KKZ judge that two — "Judicial Independence" and "Crime" — are relevant for their composite indicator "Rule of Law". So KKZ calculate the arithmetic average of these two indicators and call this number their "World Markets Online" source, which they then combine, as explained in the stages below, with 23 other sources to produce their composite indicator "Rule of Law". This particular source covers as many countries as are covered by both the "Judicial Independence" and "Crime" indicators supplied by the World Markets Research Centre.

In a second stage, KKZ apply a formula to each source to be used in the construction of the particular composite indicator in order to determine whether the source covers a large enough number of countries in different income categories and regions to qualify as a "*representative*" source. For the composite indicator "Rule of Law", out of the 24 sources used in 2004, ten qualify as "representative" (of which nine are "Expert Assessments" and seven are produced for commercial purposes).

In a third stage, KKZ aggregate the "representative" sources into a preliminary composite indicator on Rule of Law. However, contrary to the first aggregation procedure, in stage one (which gave the same weight to each existing indicator in producing a single number for each of the 24 sources), the aggregation procedure used in this stage is *not* a simple average. Instead, the different "representative" sources are *weighted according to the strength of their correlation with one another*. The more closely a source correlates with other sources, in other words, the smaller is its estimated "error variance" and the greater is its weight in the calculation of the composite indicator. (See both Box 4.2 and Appendix I for greater detail.)

Underlying this stage-three procedure is a crucial — and unrealistic — assumption: that different sources' errors are *uncorrelated* with one another, so that a high degree of correlation between the numbers shown by some sources is *not* a reflection of a correlation of these sources' measurement errors, but instead a reflection of their greater accuracy, compared to less closely correlated sources, in terms of the underlying reality of governance they are designed to reflect. One of our greatest concerns is precisely the extent to which users of governance indicators fail to understand the significance of this assumption, to which we return below.

ISBN 92-64-02685-1 © OECD 2006

Box 4.2. **Weighing the Sources**
(stage 3 in the construction of a KKZ composite indicator)

In their aggregation of the numbers from different sources to construct each of their six composite indicators (the numbers from the 24 sources used to construct the "Rule of Law" indicator for example) KKZ do not give the same weight to the numbers from each source. Rather, the weight assigned to the numbers from each source is inversely proportional to its error variance relative to the other sources used to construct the indicator, so that the more closely the numbers from one source correlate with those of other sources, the lower is the source's perceived error variance and the higher is the weight KKZ give to its numbers in the aggregation process they use to construct the composite indicator.

The logic of this weighting procedure is based on the assumption that errors in the numbers of one source are both independent across countries (for the same source) and independent of the errors in the numbers from the other sources used to construct the same composite indicator. To illustrate, any error in the Heritage Foundation's assessment of China's quality of governance is assumed to be independent both of any error in the Heritage Foundation's assessment of the quality of governance in other countries and independent of other sources' errors in the assessment of the quality of governance in China.

The assumption that sources' errors are mutually independent and therefore uncorrelated across sources and countries means that any statistical correlation found among the numbers of different sources is interpreted by KKZ as due exclusively to their correlated sources' greater factual accuracy, compared to other sources used in constructing the indicator. In other words, sources that tend to be highly correlated with other sources are *assumed* to be more informative, and therefore KKZ attribute lower error variances and higher weights to them than to sources that are less closely correlated with the majority. The latter sources are given low weights and their influence on the composite indicator is therefore also low.

The way the sources are weighted is thus very sensitive to the composition of the group of sources used — and to changes in that composition. If, by way of illustration, five sources were to be aggregated in one composite indicator and four of them were very similar, the weights of the four similar sources will be overwhelming and the source that differs will have practically no weight in the construction of the composite indicator. Moreover, increasing the number of similar sources tends automatically to decrease the weight of less similar sources much *more* than proportionally, because the perceived error of the less similar source is much greater due to the fact that it differs from the newly introduced source.

ISBN 92-64-02685-1 © OECD 2006

Box 4.2 (contd.)

Since most sources do not cover all countries, the weights for each source differ from one country to another. This difference poses a problem for the comparability of countries' scores.

It is also important to note, finally, that the weight given to the numbers from a particular source, relative to those from another source, does not depend on the sample size on which the numbers of either source are based (because the error variance imputed to each source is not affected by its sample size). Thus, for example, whether an expert assessment draws on 5 or 500 experts, or a household survey interviews 100 or 10 000 households, does not affect its weight.

In a fourth stage, KKZ regress the "non-representative" sources on the preliminary composite indicator (produced in stage three) to obtain estimates of the error variances of these sources (i.e. the other 14 sources used to produce the "Rule of Law" composite indicator). Sources that correlate more strongly with those weighing most heavily in the calculation of the preliminary composite indicator will, again, have lower estimated error variances and therefore been given higher weights.

Finally, KKZ calculate new weights for all sources based on their respective error variances (obtained in stage three for the "representative" sources, in stage four for the "non-representative" sources) and aggregate them, according to their weights, to produce the composite indicator.

Three further features of the KKZ method of producing their composite indicators are worth flagging. One is that while the procedure is the same for constructing all six composite indicators, and many sources are of course used in the construction of more than one composite indicator[2], the weights assigned to each source differ from one composite indicator to another, from one country to another as well as from one year to another. The reasons for this variation are that *a)* the number and composition of sources, for a given country in a given year, are not the same for each composite indicator[3]; and *b)* the numbers of one source may be highly correlated with those of other sources used in the construction of one composite indicator for a given country in a particular year, but not with those of the sources used for another composite indicator for that country (or for another country, or for another year).

ISBN 92-64-02685-1 © OECD 2006

Second, and particularly important for our purposes, is the fact that KKZ construct each of their composite indicators in such a way that *the average value of the indicator* across all countries, worldwide, *is always zero* and its *standard deviation is always one*[4]. This technical feature of the KKZ indicators signals that their scale is largely arbitrary and that they cannot reliably be used for monitoring changes in levels of governance over time, whether globally, in individual countries or among specific (e.g. regional) groups of countries[5]. Unfortunately, this feature tends widely to be ignored by users.

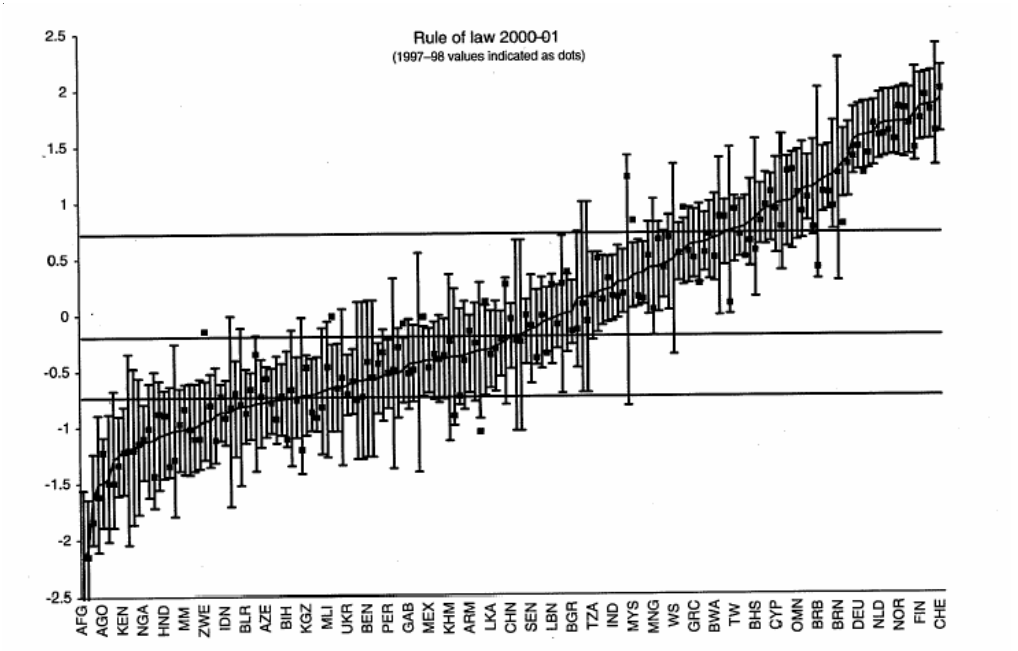
Third, for each country covered by a given composite indicator, KKZ provide both a precise numerical score (point estimate) and a "confidence interval" around that score which, according to their estimate, has a 90 per cent probability of containing the country's "true" (as opposed to estimated) score (see Box 4.3). To illustrate, Figure 4.1 shows country scores and their corresponding confidence intervals for the composite indicator "Rule of Law" in 2000[6] and in 2004.

The size of the confidence intervals reflects KKZ's estimate of the size of the measurement error embodied in the point estimates for each country in a given year. The smaller KKZ believe the measurement error to be for a given composite indicator, the smaller is the size of the confidence intervals they show for countries' scores on that indicator. The 90 per cent probability that KKZ's confidence intervals contain countries' "true" levels of governance depends, of course, on the validity of the assumptions KKZ make in constructing the confidence intervals (notably including the one that sources' measurement errors are uncorrelated across sources and countries).

While one may legitimately question the reliability of the assumptions reflected in the size of the confidence intervals, it is equally important to emphasise that measurement errors are by no means unique to the KKZ indicators. On the contrary, in producing such confidence intervals, Kaufmann and his co-authors are among the few producers of governance indicators who try to raise awareness among potential users of governance indicators of the serious measurement problems inevitably associated with such indicators. They draw attention to the fact that users should not try to distinguish between the levels of governance in countries whose confidence intervals overlap, even if their scores appear quite different from each other (and all the more so if countries' scores appear close to each other). In doing so, KKZ seek to draw users' attention to the importance of measurement errors, and of using governance indicators with proper caution.

ISBN 92-64-02685-1 © OECD 2006

Figure 4.1. **Imprecision of Governance Estimates**[a]
Figure 4.1A

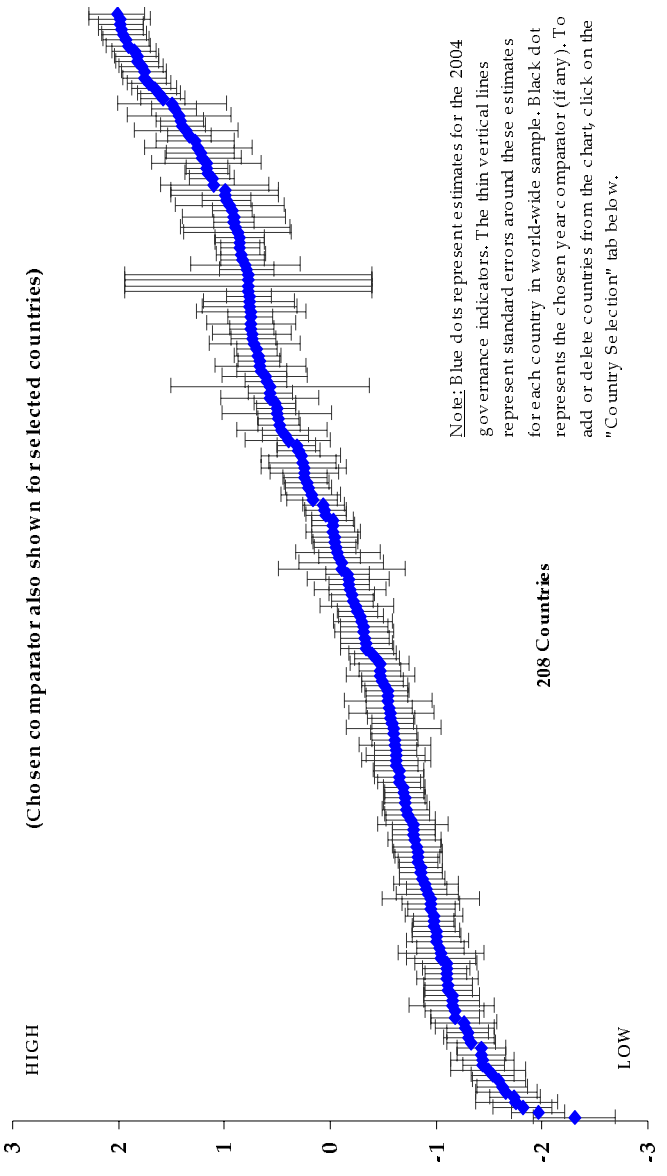


Note: *a)* This graph shows estimates of the indicated dimension of governance (on the vertical axis) for all countries for which data are available (on the horizontal axis) for 2000-01. The vertical bars show the statistically likely range of values for each country, with the midpoint of each bar corresponding to the best single estimate. The length of these bars varies with the amount of information available for each country and the extent to which information from different sources corresponds with each other. Estimates of governance for 1997-98 are indicated as dots. Selected countries are indicated on the horizontal axis. As emphasised in the text, the ranking of countries along the horizontal axis is subject to significant margins of error, and this ordering in no way reflects the official view of the World Bank, its executive directors, or the countries they represent.

Source: authors' calculations, as described in Kaufmann *et al.* (2002).

ISBN 92-64-02685-1 © OECD 2006

Figure 4.1B



Source: "Governance Matters IV: Governance Indicators for 1996-2004 " by Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi, 2005. Disclaimer: The governance indicators presented here reflect the statistical compilation of responses on the quality of governance given by a large number of enterprise, citizen and expert survey respondents in industrial and developing countries, as reported by a number of survey institutes, think tanks, non-governmental organisations and international organisations. The aggregate indicators in no way reflect the official position of the World Bank, its executive directors, or the countries they represent. As discussed in detail in the accompanying papers, countries' relative positions on these indicators are subject to margins of error that are clearly indicated. Consequently, precise country rankings should not be inferred from this data. Photographed from:"Governance Indicators: 1996-2004—Global Comparative Charts", http://www.worldbank.org/wbi/governance/govdata/excelgraphs.html.

ISBN 92-64-02685-1 © OECD 2006

Box 4.3. **Confidence Intervals and Measurement Errors**

As Figure 4.1 illustrates for "Rule of Law" in 2000 and 2004, an important feature of the KKZ indicators is that, for each country covered by a given composite indicator, the authors provide both a numerical score (point estimate) and a "confidence interval" around that score which is estimated to include the "true" (as opposed to estimated) score with a 90 per cent probability.

KKZ define the confidence interval as the country's score plus and minus 1.64 times its standard error*. Based, again, on the assumption that measurement errors are *not* correlated across sources, KKZ's estimation of the standard error of a country's score depends in part on *a)* the number of sources used to generate the composite indicator for the country; and *b)* the estimated accuracy of each of these sources. The greater the number of sources they use to generate the composite indicator for the country and the more closely these sources are correlated with each other (thus the greater their assumed accuracy), the lower is the standard error KKZ calculate for the country's score — and the smaller (better) will be the country's confidence interval.

To illustrate, Kaufmann *et al.* (2003) explain that the confidence interval for a country with only one source** will be about twice as large as the confidence interval for a country with seven sources (the median number of sources). The fact that the standard errors and corresponding confidence intervals for countries shown in Figure 4.1 for 2004 are smaller than those shown for 2000 can thus be the result of *a)* an increase between 2000 and 2004 in the number of sources KKZ use to produce the indicator and/or *b)* a higher degree of correlation among the sources (and correspondingly lower standard errors) in 2004 than in 2000. This latter cause may in turn be the result of either *i)* an increase in the degree of accuracy of the sources; or, on the contrary, *ii)* an increase in the degree of correlation of measurement errors among sources — i.e. an increase in the degree of violation of a core assumption underlying the method of construction of the confidence intervals.

Because users — including us — do not have access to the all the sources KKZ use to produce their composite indicators, we cannot determine whether factor *a)* or *b)* appears to be the major driver behind the "shrinking" of the confidence intervals between 2000 and 2004. It is nevertheless worth emphasizing that in the absence of correlation of their sources' measurement errors, a higher correlation among sources means an increase in the sources' accuracy as KKZ assume***; but if, on the contrary, an increase in correlation among sources is actually driven by increasing correlation in their measurement errors (due, say, to experts' increasing mutual influence on one another's perceptions), then the true size of the confidence intervals for countries' scores should not *shrink*, as shown, but rather *increase*.

ISBN 92-64-02685-1 © OECD 2006

Box 4.3 (contd.)

Also important to note, finally, is the fact that the sample size of an underlying source does not affect KKZ's estimate of the source's degree of accuracy, and therefore does not affect their estimate of the standard error of a country's score on a composite indicator using that source.

* According to Kaufmann *et al.* (2005*b*), the standard error reflects the measurement error embodied in the point estimates for each country in a given year. Kaufmann *et al*. refer interchangeably to "standard error" and "margins of error".

**Between 5.7 and 8.1 per cent of all countries covered by a composite indicator (depending on the indicator) have only one source. Most are small countries such as American Samoa, Anguilla, Aruba, French Guiana, Guam, Macao, Netherlands Antilles, Reunion and the Virgin Islands. The average number of sources used per country ranges between 7 and 9 (see footnote 2) and the standard deviation of that number between 3 and 5.

*** This is true under the assumption that all sources measure the same concept of governance (e.g. measure the same concept of rule of law). If the sources measure different aspects of governance, high correlation can no longer be interpreted to indicate high accuracy. The reason is that we do not only want to extract from the sources the information they have in common, but more importantly we want to extract the *different* information each source brings to bear on the quality of governance (e.g. on the quality of rule of law). We would therefore be interested precisely in the non-correlated variation that is discarded by KKZ as measurement error.

Concerns

These technical features of the KKZ indicators (indicators which, it bears repeating, are probably the most carefully constructed), together with our earlier observations about how they are widely used, point up four major reasons for concern about their *misuse*: the likelihood of correlation of errors among sources; lack of comparability over time; sample bias; and insufficient transparency.

(i) Violation of the assumption of non-correlation of sources' errors

A core assumption of the aggregation procedure which KKZ use to construct their composite indicators is one the authors themselves recognise as unrealistic: that their sources' errors are uncorrelated. At least four types of reasons, often mutually reinforcing, and for which there is considerable evidence, explain why their sources' errors often *are* in fact correlated:

ISBN 92-64-02685-1 © OECD 2006

1) Experts who supply perception data used in one KKZ source are often informed of and influenced by the assessments of experts supplying such data for other KKZ sources (sources produced both by the same organisation and by other organisations).

2) Experts who supply perception data for diverse KKZ sources are often informed of and influenced by perceptions and assessments from the same third parties (non-KKZ sources).

3) Perceptions used as inputs for KKZ "governance" indicators are often influenced, significantly and in similar ways, both by crises (financial and/or political) and by perceived changes or longer-term trends in a country's *economic* performance, FDI inflows, etc.

4) Because the interpretation of survey questions is context- and culture-specific, perception errors of different sources that rely on respondents from the same country or culture are likely to be correlated.

Of the many examples one could cite to illustrate the first two reasons (please refer also to Box 4.1), four suffice: *i)* The World Bank (2004*c*) advises its staff responsible for producing CPIAs (which serve KKZ as a source) to use, among others, the KKZ indicators and some of their sources (e.g. ICRG, the Heritage Foundation's Index of Economic Freedom). *ii)* Freedom House supplies indicators that KKZ use in constructing three different sources. *iii)* Amnesty International and the US State Department supply human rights data used by both the University of North Carolina's "Political Terror Scale" and the University of Binghamton's "Cingranelli & Richards Human Rights Database" which KKZ use as different sources. *iv)* The Economist Intelligence Unit, which is one of KKZ's main sources, uses a version of Transparency International's CPI "cleansed" of the EIU's original data as a benchmark for its own ratings, and the CPI uses practically the same sources as the KKZ indicator "Control of Corruption" (Galtung, 2005).

Considerable evidence also supports the causes for concern cited under reason 3, i.e. that many experts who provide assessments used as inputs for KKZ composite indicators are significantly influenced both by longer-term economic trends and short-term crises in the countries whose "governance" qualities they are supposed to be assessing (Knack, 2002). The same hysteresis and herd effects that plague financial actors and risk-rating agencies, noted earlier, also often lead the experts and business managers whose perceptions are at the base of many KKZ indicators to be too optimistic or too pessimistic — and to violate the assumption of non-correlation among their perception errors. An illustration is the behaviour of both the CPI and the ICRG corruption indicators, both of which showed corruption in Indonesia to be falling until

ISBN 92-64-02685-1 © OECD 2006

the 1997 Asian financial crisis, and rising after the crisis — whereas recent analysis finds that actual corruption in Indonesia did not rise after the crisis (Thompson and Shah, 2004).

Does such correlation of errors matter? Unfortunately, significant violations of the assumption of non-correlation of sources' errors have significant negative implications for the reliability of the KKZ indicators. Correlation of errors among sources means that each additional source used to produce a given composite indicator actually contributes less additional information to the construction of the composite indicator than the authors assume, and the margins of error are greater than they calculate[7]. The true 90 per cent "confidence interval" for individual countries' scores is therefore bigger than those they calculate (or, put differently, the interval they show for each country offers less than a 90 per cent level of confidence) and, more importantly, the statistical significance and reliability of cross-country comparisons among countries' scores is (even) less reliable than the authors estimate.

To their credit, KKZ explore the likely significance of departures from their assumption of zero correlation among their sources' errors. They do so by raising the assumed level of correlation among sources' errors from 0 to 0.5 and examining its effect on a sample of indicators and countries (Kaufmann *et al.*, 1999*a*). They report that the average standard error[8] *doubles* for "Rule of Law" (from 0.33 to 0.66). The effect, in other words, is literally to double the size of confidence intervals, and significantly weaken users' ability meaningfully to compare countries using this indicator.

Yet KKZ also report that the effect of raising the assumed level of correlation of sources' errors from 0 to 0.5 varies considerably from one composite indicator to another. Thus, for "Government Effectiveness", they find a smaller impact of raising the assumed level of correlation among sources' errors from 0 to 0.5: it raises the average size of the standard error from 0.32 to 0.35[9]. While reassuring in the sense that the reliability of other composite KKZ indicators may not be as seriously weakened as that of "Rule of Law" by the fact that errors among sources used in their construction tend to violate the assumption of non-correlation, this finding points up another concern to which we return below: a relative lack of *transparency* of the indicators[10].

(ii) Lack of comparability over time

The KKZ composite indicators do not allow for a reliable comparison of levels of governance over time — whether for the purpose of monitoring change within a country, comparing apparent progress or deterioration in one country

ISBN 92-64-02685-1 © OECD 2006

with others, or discerning regional or global trends — as noted earlier. The KKZ indicators are therefore not appropriately used for rewarding improvements or punishing deterioration in governance over time. Yet this is precisely how many users apply them in practice.

Recognising the potential seriousness of the problem created by the extent to which users of the KKZ indicators tend to draw inferences and make decisions based on a comparison of countries' scores over time, Kaufmann *et al.* (2005*a*) propose a "rule of thumb" for using their indicators which seeks to address this problem[11]. The rule of thumb they propose is to use their indicators to identify countries whose change in scores over time is large enough, relative to the size of the scores' confidence intervals, to ensure that the change in scores is likely to reflect real change and not merely constitute a statistical artefact. The rule of thumb is that if, in comparing a country's score in a given indicator in two different years, one finds that the confidence intervals for the two different scores do not overlap, one can be reasonably confident that change has actually occurred in the country during that period, as regards the aspect of governance to which the indicator refers, in the direction (improvement or deterioration) shown by the change in scores[12].

Applying this rule of thumb to their indicators for 1996 and 2004, KKZ find that of the approximately 200 countries covered by their indicators, about a quarter (56 countries) had scores in at least one indicator that changed enough over the eight-year period, compared to their confidence intervals, to suggest actual change had occurred in the quality of governance in those countries[13]. For each indicator, however, only a small number of countries' scores showed enough change to satisfy the rule of thumb: between about 5 and 8 per cent of all countries covered, depending on the indicator[14]. Thus, of the 56 countries whose change in scores satisfied the rule of thumb suggesting actual change, most did so for only one or two of the six KKZ indicators (37 countries showed significant change in only one indicator, 11 countries showed such change in two indicators); in only eight countries did scores change significantly in three or more indicators — and in three of these the scores showed both improvement and deterioration.

These figures highlight the limited number of countries, out of the approximately 200 covered by the KKZ indicators, for which, according to KKZ's rule of thumb, users can reasonably infer that change actually occurred between 1996 and 2004. The very limited extent to which such change in countries' scores in one indicator is accompanied by such change in their scores in other indicators further highlights the danger of using the KKZ indicators

ISBN 92-64-02685-1 © OECD 2006

to draw broader inferences about the nature, the causes, or even the *existence* of change in the quality of "governance" — be it within a country or in cross-country comparisons — over time. Many users seem to ignore these limitations.

(iii) Sample bias

While the list of sources used to produce the KKZ indicators (see Box 4.1) appears to be reasonably diverse and representative of different stakeholders — ranging from population surveys to expert assessments and enterprise surveys — the aggregation procedure used to calculate the composite indicators assigns less weight to sources that differ from the majority, as explained previously. The result of this procedure is effectively to give much more weight to expert assessments and enterprise surveys than to population surveys — to the point that population surveys carry practically no weight in the composite indicators[15].

The "Voice & Accountability" indicator illustrates: sources with the highest weights are Freedom House's "Nations in Transition" (0.39) and "Freedom in the World" (0.12) and the Economist Intelligence Unit's "Country Risk Service" (0.18), whereas sources based on the population surveys Afrobarometer, Latinobarometer and Gallup International's "Voice of the People" carry practically no weight (0.01)[16].

The reason for these marked differences in weights is precisely the differences in the perceptions of the quality of local governance reported by the different groups picked up in the different sources. These differences are confirmed by recent "mirror surveys" in francophone Africa which show that the assessments of the quality of local governance given by "experts" differ substantially from those given by respondents to local population surveys[17]. For the KKZ indicators to reflect the diversity of perceptions provided by the population surveys, relative to those provided by expert assessments and business surveys, the population surveys would have to be given more than negligible weights in the construction of the indicators. But the opposite is the case[18].

This sample bias in favour of business-oriented perceptions is not unique to the KKZ indicators. Transparency International's Corruption Perceptions Index (CPI), which draws on many of the same sources as KKZ's "Control of Corruption" indicator, is characterised by a similar bias. As one of Transparency's founders and former head of research has noted, "Of the 17 different institutions providing data for the CPI since 1998, only two do not

ISBN 92-64-02685-1 © OECD 2006

have a private-sector bias… This homogeneity…generates a serious sample bias, which is a genuine failing of the CPI. The sample is not only private sector oriented, it is also overwhelmingly male and economically well off. Effectively, this means that this most influential of indices ignores the experiences and perspectives of most women, and of the poor and disenfranchised. It also means that the interests of 'unofficial businesses', which employ the overwhelming majority of the population in poor countries, are ignored" (Galtung, 2005).

This bias in favour of business surveys and expert assessments[19] (both of which tend to be oriented toward the interests and perceptions of relatively large formal-sector investors and corporations), which in the KKZ indicators is exacerbated by their aggregation procedure, is problematic for at least three reasons.

One, in the words of the World Bank's *2005 World Development Report,* is that "although societies benefit greatly from the activities of firms, the preferences of firms don't fully match those of society. This tension is most evident in taxation and regulation. Most firms complain about taxes, but taxes finance public services that benefit the investment climate and other social goals. Many other firms would also prefer to comply with fewer regulations, but sound regulations address market failures and can therefore improve the investment climate and protect other social interests" (World Bank, 2004*a*). Yet, for example, the KKZ indicator on "Regulatory Quality" gives considerably greater weight to sources that reflect individual business managers' views on both labour and environmental regulations as harmful "for the growth of your business" and "for competitiveness" than it does to sources, notably population surveys, that reflect others' markedly different perceptions on the role of such regulations[20]. The danger is that users of this indicator could *unwittingly* be influenced in their decisions by a hidden bias against labour and/or environmental regulations, for example.

Second, there is a danger of circular reasoning. Investors' herd behaviour and the number of major financial crises in developing countries over the last decade suggest that investors' confidence levels are not always good indicators of actual governance realities in a country. Investors themselves increasingly see this problem. As a US-based international business spokesperson[21] interviewed for this study phrased the problem with today's governance indicators: "We are basically talking to ourselves. We need indicators that give us a more objective understanding of the actual conditions of governance in developing and emerging-market economies".

ISBN 92-64-02685-1 © OECD 2006

Third, whether in good faith (as is often the case) or not, governments in developing countries often resent a perceived private-sector and "western" bias in the indicators. An important consequence, reported by a major OECD official donor agency, is that the use of such governance indicators in the agency's discussions with recipient governments often left the latter feeling accused without providing them with any meaningful guidance on the causes of their governance shortcomings or on how effectively to overcome those shortcomings. The result was often to polarise discussions, leave recipient governments feeling alienated, and weaken the very basis for collaboration between the agency and governments it was intending to help.

A good illustration of such reactions is provided by the Secretary General of the Association of Caribbean States, Norman Girvan, in his comments on the use of governance indicators (including the KKZ indicators, those of Freedom House, and others) in the 2002 UNDP *Human Development Report*: "[While] most of the OECD countries get perfect scores in the indicators…[w]ell-known phenomena in the perversion of these [countries'] political systems by powerful financial groups and vested interests are inexplicably omitted. This would require examining factors such as the transparency of campaign financing, limits on party political funding and the impact of vested interests on economic decision-making, for example in the area of trade policy. The developed countries also score highly in Press Freedom — notwithstanding the domination of the media by huge conglomerates and the slanted nature of much press coverage of North-South issues — and in Legal Impartiality — although it is well known that the wealthy are much better equipped to utilise the justice system than the poor and ethnic minorities" (Girvan, 2002).

Much to his credit, and undoubtedly in response to such concerns, in a separate article Kaufmann gives explicit attention to the governance deficiencies that widely afflict OECD and developing countries alike. The problem, he explains, is that "the undue emphasis on narrow legalism has obscured more subtle yet costly manifestations of misgovernance, which afflict not only poor but rich countries… In particular, where the 'rules of the game' have been captured by elites, frequently ignored manifestations of so-called 'legal corruption' may be more prevalent than illegal forms, such as outright bribery, which are the usual focus of attention" in governance indicators (Kaufmann, 2004). He has thus developed a "legal corruption" indicator that focuses on such problems of corruption and state capture in rich and poor countries alike — which nevertheless remains separate from the KKZ indicators (Kaufmann and Vicente, 2005).

ISBN 92-64-02685-1 © OECD 2006

(iv) Lack of transparency

It is important for users of governance indicators to be able to understand and even, if they wish, to be able to replicate the calculations by which the indicators they use are obtained. Few governance indicators have this quality of transparency. Unfortunately the KKZ indicators are no exception, for at least four reasons.

First, the large number and diversity of indicators produced by others that are incorporated into the KKZ indicators makes it very difficult to understand why a country gets a particular score on a given composite indicator. This difficulty is all the greater because some of the indicators used as inputs are themselves very broad and imprecise or lacking in transparency. For example, existing indicators used to construct the composite indicator "Voice & Accountability" range from "imprisonment" to "freedom of the press" to "vested interests".

Second, there is no list of criteria that each source uses to determine its country rankings, so of course there is no such list for the composite KKZ indicators. The judgement of what constitutes good governance depends entirely on each source, and none reveals its exact criteria. It is therefore impossible to explain why a country obtains a particular score, or what concretely it should do to improve that score.

Third, some of the data from the sources are not accessible and others are difficult or expensive to access. For example, the CPIAs remained unaccessible outside the World Bank until very recently, and access to the "Global Risk" data supplied by Global Insight costs $12 700.

Fourth, the fact that some sources comprise existing indicators that refer either directly or indirectly to economic and business growth or competitiveness creates the risk, noted earlier, of circular reasoning. For example, some of the existing indicators used to construct the KKZ composite indicator "Regulatory Quality" are based on respondents' assessments of whether local regulations (notably labour and environmental regulations) constitute obstacles to growth or investment. Studies that find a link between "Regulatory Quality" and growth or investment are therefore subject to circular reasoning.

It should also be clear that, difficult as it is to interpret the meaning of countries' scores from individual KKZ composite indicators, studies that combine several or all of the KKZ indicators into a single "governance" indicator — which KKZ never do — produce results that in our view (see the section on analysts and academics in Chapter 3) are essentially meaningless.

ISBN 92-64-02685-1 © OECD 2006

Notes

1. As shown in Box 4.1, a KKZ "source" commonly comprises several indicators produced by one organisation (e.g. the various ICRG indicators produced by the Political Risk Service), but a single source may also comprise indicators produced jointly by two or more organisations (e.g. the Business Environment and Enterprise Performance Survey — BEEPS — produced jointly by the World Bank and the EBRD); a single organisation may also produce more than one source (e.g. the World Bank produces both the CPIA and BEEPS, Freedom House produces both "Freedom in the World" and "Nations in Transit" which rely on different survey data and KKZ use as separate sources).

2. Of the 37 sources used to produce all six composite indicators in 2004 (see Box 4.1), 19 were used for producing the "Voice and Accountability" indicator, 13 for producing the "Political Stability" indicator, 19 for "Government Effectiveness", 15 for "Regulatory Quality", 24 for "Rule of Law" and 22 for "Control of Corruption". The average number of sources for each country covered by a composite indicator in 2004 was between 7 (for both "Government Effectiveness" and "Regulatory Quality" composite indicators) and 9 (for "Rule of Law").

3. The KKZ indicators may therefore measure slightly different things in different countries (Van de Walle, 2006).

4. Practically all countries' scores thus lie between -2.5 and +2.5, with higher scores signalling better governance.

5. Kaufmann *et al.* (2005*a*), explain that there is no evidence for trends in global averages in governance in their underlying sources over the period they look at. One may therefore be tempted to conclude that their indicators are comparable over time. We do not share this belief, however, because *a)* there is no evidence to the contrary either, i.e. there is no empirical basis on which to conclude that there are no trends or other important types of variation in the global quality of governance over time; *b)* not only would global averages need to be constant, so too would standard deviations need to be constant, to allow for reliable comparisons of levels of governance over time on the basis of the KKZ indicators, yet there is also no reason to believe that the standard deviations are constant,

ISBN 92-64-02685-1 © OECD 2006

and *c)* many underlying sources used to construct the KKZ indicators do not allow for reliable comparisons of levels of governance over time. Thus, to understand the practical consequences of using KKZ composite indicators to compare levels of governance over time (whether within a country or between countries), suppose the *actual* ("true") level of governance in a country remains *unchanged* over time. Its score on any of the KKZ indicators is nevertheless likely to change for one or more of three reasons: *i)* change occurs in the score of one or more of the other 200-odd countries covered by the indicator (e.g. other countries' improvement, other things equal, causes a deterioration in our country's score); *ii)* KKZ include one or more new sources in the construction of the indicator, which causes change in the weights given to all sources; and *iii)* one or more new countries are included in the country coverage of the indicator (e.g. if a new country has a better score than our country, its inclusion in the coverage of the indicator will automatically, other things equal, lower the score of our country).

6. As published in Kaufmann *et al.* (2002). This figure was revised in Kaufmann *et al.* (2003, 2005*a*).

7. The violation of the non-correlation assumption also probably leads to inconsistent estimators (see Appendix I).

8. According to Kaufmann *et al.* (2005*a*), the standard error reflects the measurement error embodied in the point estimates for each country in a given year (see Box 4.3). Kaufmann *et al.* (2005*a*) refers interchangeably to "standard error" and "margins of error".

9. This calculation is again based on the assumption that errors are equally correlated among sources. However, we suspect that the correlation of errors varies according to the type of source: Errors in household surveys are likely to be correlated with errors in other household surveys, errors in expert assessments with those in other expert assessments, those in business surveys with errors in other business surveys. Without access to all the underlying sources, we cannot assess ourselves how high the standard errors would be when assuming different levels of correlation of errors according to the type of source.

10. As noted in footnote 9 of Chapter 2, the World Bank Institute has reportedly decided to begin disclosing countries' scores on their underlying sources.

11. Kaufmann *et al.* (2005*a*) construct, in addition to their original "static" model, a "dynamic" model designed to allow comparisons of countries' governance scores between 1996 and 2004. While they do not publish countries' scores based on their "dynamic" model, they report that the list of countries designated by the model as countries whose quality of governance changed significantly over the eight-year interval is largely the same as that which results from applying their "rule of thumb" to their original, "static", model.

ISBN 92-64-02685-1 © OECD 2006

12. Though apparently conservative in its reliance on supposed 90 per cent confidence intervals, KKZ's rule of thumb of course ignores the weakening of these intervals caused by the violation of the assumption of non-correlation in sources' errors at a given point in time, as just discussed. This effect is distinct from and comes in addition to the effects of correlation in errors *over time*, and in *unobserved governance* over time, as analysed elsewhere in Kaufmann *et al.*, 2005*a* (Section 3).

13. In 24 of these countries the change in scores pointed to an improvement, in 29 countries it pointed to a deterioration, and in three countries it pointed to both — i.e. significant change appeared in the scores of three countries in several indicators, of which at least one showed improvement and another showed deterioration. Specifically, *Indonesia* showed an improved score in Voice & Accountability, and deterioration in Political Stability, Regulatory Quality, and Rule of Law; *Latvia* showed improvement in Government Effectiveness and Control of Corruption, and deterioration in Political Stability; *Sierra Leone* showed improvement in Voice & Accountability and Political Stability, and deterioration in Government Effectiveness.

14. The numbers range from seven countries in the case of "Government Effectiveness" (three countries' scores show likely improvement, four likely deterioration) to 22 countries in the case of "Political Stability" (six countries' scores showing likely improvement, 16 likely deterioration).

15. Even if we took a simple average of all sources instead of weighing them according to their degree of correlation, expert assessments and business surveys would dominate the overall result because there are few cross-country-comparable household surveys.

16. The weights reported here are from Kaufman *et al*. (2003) since no new weights were published in Kaufmann *et al.* (2005*a*). They are weights that would be assigned to each source for a hypothetical country appearing in all sources used to produce the indicator "Voice & Accountability". If a country appears in a subset of sources, the weights applied would be proportionate to the ones reported for those sources (Kaufmann *et al.*, 2003).

17. DIAL (2005, 2006).

18. It would be useful to run tests to answer the following questions (which we nevertheless cannot do ourselves because we do not have access to all the data): Are there systematic biases according to the type of source? Are these biases the same for all groups of countries (rich and poor, in different regions)? Are the patterns of correlation the same for all countries, or are there sources which are more highly correlated for certain countries but not for others? Interpreting the independent variation in the sources (i.e. the part that is not correlated) as "error" as do Kaufmann *et al*. may actually lead to losses of very relevant information (as noted also in the third footnote in Box 4.3).

ISBN 92-64-02685-1 © OECD 2006

19. There are few household-survey studies, or data, that are comparable across countries, whereas there is a large number of expert assessments and business surveys with large country coverages. Many composite indicators, such as the CPI and the KKZ indicators, are therefore dominated by data from experts' assessments and business surveys — not because of a deliberate choice by the producers of the composite indicators, but because of the paucity of internationally comparable household-survey data relative to the abundance of internationally comparable data from experts' assessments and business surveys. Unfortunately, the data-aggregation methods used to produce the composite indicators fail to account for this sample bias.

20. BEEPS reports business managers' answer to the question, "How problematic are labour regulations for the growth of your business" and the World Economic Forum asks them in its Global Competitiveness Survey to give their opinion on the statement, "Environmental regulations hurt competitiveness" while Global Insight focuses on "regulatory burdens…that reduce aggregate investment".

21. John D. Sullivan, Executive Director, Center for International Private Enterprise, Washington D.C.

ISBN 92-64-02685-1 © OECD 2006

Chapter 5

Governance and Growth

Summary

Current governance indicators are highly positively correlated with measures of current national per capita income. The challenge is to identify the direction(s) of causality in the relationship between the quality of governance and the level of income in a country. We find that beyond the limitations inherent in the construction of most governance indicators used for such analysis, research based on those indicators often produces results that are very sensitive to changes in the econometric model used — the variables in the model and its underlying assumptions — as shown in this chapter.

Our objective in this chapter is to show that, in addition to the limitations of governance indicators *per se*, research based on those indicators often produces results that are very sensitive to changes in the econometric model used — the variables in the model and its underlying assumptions. To illustrate we continue our focus on the KKZ indicators and examine the very important study by Kaufmann and Kraay (2002*a*) of the causal relationships between economic growth and the quality of governance. The authors' remarkable conclusion, suggested by their title "Growth without Governance", is that while better governance tends clearly to promote economic growth, growth *per se* does *not* tend to promote better governance.

Indeed, though remarkable, Kaufmann and Kraay's finding that stronger long-term economic growth tends on average to have negative feedback effects on the quality of governance in countries around the world is not inconsistent with the findings of other authors. One of the most noteworthy is Mancur Olson (Olson, 1965, 1982), whose analysis of the tendency for powerful vested-

ISBN 92-64-02685-1 © OECD 2006

interest groups to build up and increasingly weaken key aspects of governance in a country over time is brought to mind (although Olson also noted that the actions of such groups, if left unchecked, would eventually slow growth as well). Also quite consistent with Kaufmann and Kraay's finding is the ample research showing strong negative relationships between raw-materials export rents and the quality of governance in some major oil- and minerals-exporting countries (e.g. Leite and Weidmann, 1999).

Many other studies argue the opposite. Some make the point that countries whose governance systems, today or in the past, would probably receive low scores by the standards of the "Rule of Law" indicator in today's context — i.e. countries with poor governance — have achieved strong long-term economic growth. Among the examples these studies cite are China since 1978, the dictatorships that were the Asian "tigers" in the 1970s and 1980s, the economic "miracle" of military-ruled Brazil in the 1960s and 1970s, France during its *trente glorieuses* after the Second World War, and Switzerland and the United States in the 19th century (e.g. Meisel, 2004; Chang, 2002). Many authors have also argued — along the lines of "modernisation" theory — that *a)* democracy is a luxury that poor countries can afford only after considerable growth has been achieved; *b)* dictatorship is often needed to achieve such growth (because, in the words of one author, "the more democratic a government…the greater the diversion of resources from investment to consumption" (Galenson, 1959); and *c)* long-term growth will lead to the emergence of democracy — if not to ever-rising scores on the KKZ "Rule of Law" indicator (e.g. Lipset, 1959; Huntington and Nelson, 1976). Still other authors, of course, argue the converse: citing experiences as varied as those of India in recent years and the United States, they see well-functioning democracy as most reliably conducive to long-term growth because it encourages long-term investment, by safe-guarding property rights, and favours efficient policies in the long run by doing a better job of holding politicians accountable.

Kaufmann and Kraay's Model

Kaufmann and Kraay (2002*a*) start from the observation that countries' scores in all six KKZ indicators tend to be highly positively correlated with their current per capita GDP levels. The challenge is to identify the direction(s) of causality, if any, between higher governance scores (better governance) and higher per capita GDP (economic growth). Is it countries' higher levels of per

ISBN 92-64-02685-1 © OECD 2006

capita GDP that tend statistically to "explain" their better governance scores? Is it, on the contrary, countries' higher governance scores that tend to "explain" their higher per capita GDP levels? Or is it both, implying a virtuous circle of mutually reinforcing growth and governance? Or is it neither, if the correlation is due to some common but external cause? The authors proceed as follows.

They build a model with three variables, of which one is exogenous (historical settler mortality rates) and two are to be explained (endogenous) for each country. The endogenous variables are a country's current per capita GDP and its score on the KKZ indicator for "Rule of Law" in 2000. The model depicts two relationships (in the form of simultaneous equations) according to which: *i)* a country's score on "Rule of Law" explains its current per capita GDP; and *ii)* a country's current per capita GDP and its historical rate of settler mortality explain its "Rule of Law" score[1].

Using data on settler mortality, current per capita GDP and "Rule of Law" scores, the authors estimate the first equation (in which governance scores determine per capita GDP) using two different methods of calculation: the "ordinary least squares (OLS)" method and the "instrumental variables (IV)" method[2]. They use historical data on settler mortality rates as an instrumental variable for "Rule of Law"[3]. While they find a positive coefficient for the impact of "Rule of Law" with both methods, they find the coefficient is *much larger* when the IV method is used, compared to that found with the OLS method.

Kaufmann and Kraay proceed by a process of elimination to explain the significant difference between the results obtained by the two methods. They start with the assumption that one or more of only three types of factors could explain the difference. One is measurement error embodied in the data used to produce the "Rule of Law" governance scores. They eliminate this factor as unlikely, however, because they find the *size* of measurement error needed to explain the difference to be implausibly high[4].

A second factor that could explain the difference is a variable not included in their model (an "omitted variable") which has the property of tending to push countries' per capita GDPs and "Rule of Law" scores in opposite directions. The authors consider this explanation implausible as well, because they can think of no variable likely to have significant opposing effects on countries' growth and quality of governance.

They find a third explanation more plausible: higher per capita GDP levels have a negative feedback effect on "Rule of Law" scores, an effect that is removed by the instrumental variable[5]. They do not directly estimate this effect, however, because they cannot identify any convincing instrumental

ISBN 92-64-02685-1 © OECD 2006

variable needed for making such an estimate that has the required property. Such an instrumental variable would have to be correlated with per capita GDP, but have no effect on "Rule of Law" scores other than through its effect on per capita GDP, i.e. it would need to be uncorrelated with every other determinant of Rule of Law not included in the regression.

Kaufmann and Kraay interpret their model's equations as capturing the effects of governance on per capita GDP growth and vice versa *in the very long run*. To do so implies two key assumptions:

— Levels of per capita GDP were roughly the same across countries at the outset of the changes induced by the interactions captured in the model, i.e. in the distant past, so that today's dispersion of per capita GDP levels among countries is a reasonably good approximation of differences in long-term *growth* among countries[6].

— The quality of "Rule of Law" changes little over time in a given country, so that today's dispersion in governance scores among countries is a reasonable approximation of differences in the quality of governance in countries before substantial differences appeared in their levels of per capita GDP.

These assumptions allow the authors to conclude — based on their model and data for 153 countries on settler mortality rates in the 18th and 19th centuries, current GDP levels and current scores on the KKZ "Rule of Law" indicator — that *i)* improvements in the quality of governance in a country are likely to have a strong positive impact on long-term economic growth in that country; *ii)* higher long-term growth is likely to have a negative impact on the quality of governance in the long run; and *iii)* the negative effect of stronger growth on the quality of governance tends to be weaker than the positive effect of better governance on economic growth. It is of course the last of these — the net effect — that is probably reflected in the positive correlation observed between current GDP levels and all six KKZ governance indicators.

Our Model

With these debates and examples in mind we build on Kaufmann and Kraay's analysis to re-examine their finding of a negative effect of growth on governance, using the KKZ "Rule of Law" indicator for the latter. We proceed differently, however, in that we *a)* include more variables in the two equations, and *b)* estimate not only the first but both equations with both the OLS and IV methods.

ISBN 92-64-02685-1 © OECD 2006

Our *first* equation (see Box 5.1) expresses the average annual growth rate in per capita GDP of a country from 1970 to 2000 as a function of five variables for the country:

1) "Rule of Law" 1996 score;

2) average annual population growth rate over the period 1970-2000;

3) log of per capita GDP in 1970 (to account for convergence, i.e. the possibility that countries with higher incomes have lower growth rates);

4) geographical region (dummy);

5) average rate of investment as a share of GDP over the period 1970-2000[7].

Our *second* equation (see Box 5.1) expresses a country's 2002 "Rule of Law" score as a function of three variables:

1) the log of per capita GDP in 1970[8];

2) geographic region (dummy);

3) historic settler mortality rate.

Box 5.1. **Our Model**

(1) $G_j = \beta_0 + \beta_1 L_j + \beta_2 I_j + \beta_3 P_j + \beta_4 R_j + \beta_5 y_j + e_j$

(2) $L_j = \gamma_0 + \gamma_1 y_j + \gamma_2 R_j + \gamma_3 x_j + v_j$

G = average annual growth rate in GDP per capita

L = Rule of Law

I = Average investment rate (gross capital formation) as a percentage of GDP

P = average annual population growth rate

R = a vector of regional dummies (with β_4 a vector of coefficients)

y = log per capita GDP

x = history (proxied by settler mortality)

e = error term

v = error term

j = country

ISBN 92-64-02685-1 © OECD 2006

The motivation for our choice of variables in the two equations is threefold: *a)* we wish to account for possible bias caused by omitted variables by including more variables that may significantly affect per capita GDP growth and "Rule of Law" scores, respectively, in our two equations; *b)* we wish to distinguish the effect of income growth and income levels (standards of living) over a time horizon of the past 30 years[9]; and *c)* we wish to relax somewhat one of Kaufmann and Kraay's assumptions mentioned earlier, namely that the quality of governance in a country does not change much over time, so that its current governance score is a good proxy for the quality of its governance before substantial per capita GDP differences emerged across countries.

We estimate our first equation with OLS, then with IV, methods. Following Kaufmann and Kraay we use settler mortality as an instrument for "Rule of Law" and test the sensitivity of our results to the choice of instrument (by using the natural logarithm of the number of years of independence[10] and the historical adult illiteracy rate in 1970[11]). In line with Kaufmann and Kraay's results, we find that the coefficient for "Rule of Law" is much higher when the method of estimation is IV than when it is OLS.

We then estimate the second equation (which explains countries' "Rule of Law" scores with their per capita GDP levels) with IV in addition to OLS methods. If "Rule of Law" changes very slowly, then current "Rule of Law" may be related to "Rule of Law" in 1970. Since "Rule of Law" in 1970 is unobserved and therefore not in our equation, it will be found in the error term. For a simple OLS regression to be valid, every explanatory variable has to be uncorrelated with the error term. However, "Rule of Law" in 1970 in the error term is likely to be highly correlated with GDP levels in 1970. Based on the fact that infant mortality rates are highly correlated with per capita GDP levels, yet are unlikely to have an impact on "Rule of Law" scores other than through their correlation with per capita GDP, we use infant mortality rates per 1 000 live births in 1970 as an instrumental variable for per capita GDP levels in 1970[12].

Our findings do not confirm the conclusions of Kaufmann and Kraay (2002*a*). We find a significant *positive* coefficient for the impact of per capita GDP on "Rule of Law", suggesting in other words that living standards tend to have a positive impact on governance, as well as vice versa. The positive coefficient is in fact significantly higher when the method of estimation is IV than when it is OLS[13]. We test the sensitivity of our results to the choice of instrument by using life expectancy at birth, the log of historical per capita GDP levels in 1913 (Maddison, 2003), and the number of passenger cars per

ISBN 92-64-02685-1 © OECD 2006

1 000 inhabitants, in various combinations, as alternative instruments for per capita GDP. Our results prove to be insensitive to the choice of instruments[14]. (See Appendix II for the results of the test of the over-identifying restrictions.)

While we cannot confirm Kaufmann and Kraay's finding of a negative feedback effect of growth on the quality of governance, we do not, however, infer that growth necessarily leads on average to better governance. Rather, we feel that we should look for other possible explanations, beyond their original three. One possibility, and avenue for further research, to plausibly explain the difference between the OLS and IV estimates in both equations — theirs and ours — is *heterogeneity*[15]. A model based on panel data, which include time-series as well as cross-sectional data, could account for heterogeneity. Unfortunately, the lack of comparability over time of scores on the KKZ composite indicators, explained earlier, makes it problematic to construct such a model.

Furthermore, Kaufmann and Kraay's finding of a negative impact of growth on the quality of governance might be true for a subgroup of countries, or there might be different stages of the relationship between governance and living standards. A model that takes into account non-linearities may therefore be more appropriate than both our and Kaufmann and Kraay's linear model.

Growth and Democracy

An alternative approach to analysing the causal relationships between governance and growth worth noting is one developed in Przeworski *et al.* (2000). Their analysis relies not on *perception* data on governance, but on objective criteria for distinguishing on a yearly basis between democracies and non-democracies (and among the latter between "authoritarian" and "bureaucratic" dictatorships[16]) based on transparent, objective and accessible data for 141 countries between 1950 (or a country's year of independence) and 1990. While their choice of criteria for regime definition can of course be contested, the transparent construction of their data set makes it relatively easy to replicate their analysis. That transparency also gives clear substantive content and objective "meaning" to their data and analysis (including the limitations of the data).

ISBN 92-64-02685-1 © OECD 2006

One of the remarkable findings of Przeworski *et al.* (2000) is that while the economies of democracies and non-democracies grow on average at about the same rates, non-democracies tend to have higher *population* growth rates — women have more than twice as many children, perhaps for their "insurance value", i.e. as a hedge against greater economic uncertainty — which means that per capita growth tends to be higher in democracies. The authors also find that democracies and non-democracies tend to follow different growth *paths*:democracies make more efficient use of technology and human capital, exploit labour and women less, help people live longer, spend more on social programmes, and use a smaller labour force more productively than non-democracies. The latter depend more on capital and use a larger labour force less productively:they pay workers significantly less.

The authors also find evidence of a poverty trap. In the poorest countries, political regimes make no difference for growth, probably because states — democracies or otherwise — lack the resources necessary to make a difference. Democracies also become dictatorships more frequently in poor countries.

Regarding the impact of growth on regime type, the authors find that per capita GDP has a very significant (lagged) effect on the probability of a country's making a transition to democracy, as well as vice versa. They also find the effect of per capita GDP on *stabilising* democracies — so they don't become non-democracies — significant. Specifically, the "life expectancy" of a non-democracy (i.e. the average waiting time for a democratic transition) shrinks from 20 years at a per capita GDP level of $1 000, to 11 years at $10 000, and 6 years at $20 000[17].Growth thus helps countries to become democracies, and it also, and more powerfully, helps them to remain democracies.

ISBN 92-64-02685-1 © OECD 2006

Notes

1. A simultaneous equations model jointly explains two or more endogenous variables in which each endogenous variable can be a function of other endogenous variables as well as of exogenous variables (and an error term). An endogenous variable in a simultaneous equation model is a variable that is explained by the equations in the model (as opposed to being determined outside the model). The explanatory variable "Rule of Law score" in the first equation is explained by the second equation. The explanatory variable "per capita GDP" in the second equation is explained by the first equation. They are therefore both endogenous.

2. The advantage of the IV method over the OLS method is that it corrects for the endogeneity of an explanatory variable — in this case "Rule of Law" scores in the first equation. Contrary to the OLS method, the IV method therefore takes into account the fact that "Rule of Law" scores are themselves determined by per capita GDP levels and the other variables in the second equation.

3. Acemoglu, Johnson et Robinson (2001) were the first to use data on settler mortality rates in different countries in the 18th and 19th centuries as instruments for the quality of governance in those countries today, on the grounds that colonial powers had weak incentives to establish the institutions of good governance, and stronger incentives to establish institutions for rent extraction (i.e. poor governance institutions), in colonies where a permanent European presence had great difficulty to take root because of high settler mortality rates. As data on settler mortality are available for only 68 countries, Kaufmann and Kraay use data on tropical location (as measured by distance from the equator) and on colonial origins (as measured by the fractions of the population speaking English or a major European langue) to expand the size of the sample to 153 countries. They report that the results are similar for the restricted and the enlarged data sets.

4. The authors calculate the effect of per capita GDP on "Rule of Law" (second equation) by using the information on the OLS and IV regression in the first equation. Because this information is not sufficient to calculate the parameters in the second equation, the authors further need to identify the measurement error

ISBN 92-64-02685-1 © OECD 2006

in GDP data and governance scores, and the correlation between the error terms in both equations (see their paper for details). As these are unknown, they estimate the impact of per capita GDP on "Rule of Law" for different values of those unknowns to find out what values would allow for a positive impact of per capita GDP on "Rule of Law". They use the standard errors of "Rule of Law" (see previous chapter) as a benchmark for the measurement error in governance to judge whether measurement error would have to be implausibly high to allow for a positive impact of per capita GDP on "Rule of Law", and find those values to be implausibly high.

5. While the overall correlation between per capita GDP levels and "Rule of Law" scores is determined by the model's two equations (the effect of per capita GDP on "Rule of Law" and the effect of "Rule of Law" on per capita GDP), some of the correlation might also be due to other variables moving per capita GDP and "Rule of Law" in the same direction. The OLS method ignores the existence of the relationship in the second equation. In doing so, the OLS coefficient for the impact of "Rule of Law" on per capita GDP in the first equation reflects the overall correlation between per capita GDP and "Rule of Law" and not the impact of "Rule of Law" on per capita GDP. If the inverse effect of per capita GDP on "Rule of Law" (second equation) were to be negative, one would expect the "true" coefficient for the impact of "Rule of Law" on per capita GDP to be higher than stated by OLS. The IV method, in estimating the coefficient for the impact of "Rule of Law" on per capita GDP (first equation) takes the second equation into account. It thus comes closer to the "true" coefficient for the impact of "Rule of Law" on per capita GDP than the OLS method.

6. It is common among econometricians to interpret currect dispersion of per capita GDP levels among countries as reflecting differences in their long-term growth.

7. Annual average gross fixed capital formation from 1970 to 2000, depending on data availability (World Development Indicators, World Bank).

8. We also used per capita GDP in 1950, 1960, 1980 and 1990. We use historic per capita GDP as we *i)* suspect that per capita GDP changes can affect Rule of Law (or other institutions) only after time; and *ii)* want to measure prosperity as far as possible, in *tempore non suspecto*, i.e. not itself influenced by the variable it is supposed to explain: current Rule of Law. We use GDP levels, since we believe that current Rule of Law levels are more likely to be related to past GDP levels than to past growth rates; i.e. a causal link to Rule of Law is more plausible from absolute standards of living or prosperity levels. Ideally, we would like to have comparable-over-time Rule of Law indicators from the 1970s to today in order to be able to explain the direction in which Rule of Law evolves by income levels as well as by GDP growth. Since we do not have such time-comparable Rule of Law indicators since the 1970s, we look at the relationship between GDP and Rule of Law levels.

ISBN 92-64-02685-1 © OECD 2006

9. If changes in per capita GDP do not affect Rule of Law and vice versa within a 30-year time frame, their effects on each other in a much longer time period appear to be of little relevance for current policy making, though they may be from a historical research perspective.

10. Using data from *Online CIA Factbook*.

11. Using data from the World Bank's *World Development Indicators*.

12. One possible omitted variable in the error term that could be correlated with our instrument is democracy. In democratic societies, infant mortality rates might be lower and the regime type may influence the Rule of Law. We include Przeworski *et al.*'s regime variable for 1970 (and for 1980 and 1990 in subsequent tests) and find it to be insignificant (Przeworski *et al.*, 2000).

13. Interestingly, the more recent the period, the lower are the IV and the higher are the OLS coefficients for the impact of per capita GDP on "Rule of Law"; i.e. the coefficients start to converge. Possible explanations for such patterns are the subject of further, ongoing, work by Christiane Arndt.

14. The instruments we use in our two equations are admittedly imperfect. Settler mortality is likely to pick up differences in the physical environment (benign vs. harsh and unhealthy) as well as differences in the institutional environment or in the quality of early governance not included in the regression. And infant mortality is likely to be correlated with other determinants of the Rule of Law not included in the regression. We acknowledge that this instrumental variable approach may not identify the causality channel precisely, but we see it as a test of Kaufmann and Kraay's conclusion that growth tends *per se* to have a negative impact on governance quality (see Rigobon and Rodrik, 2004, for an alternative approach).

15. Heterogeneity refers to the fact that the relationship between growth and governance may differ significantly from one country to another in ways that our models do not capture. For example, the same improvement in the quality of Rule of Law may raise per capita GDP significantly more in one country than in another, *ceteris paribus*. Neither our model nor KKZ's model takes this difference, or "heterogeneity" (i.e. differences in coefficients and/or in variances across countries, in technical terms), into account. As the models do not account for the individual differences across countries, they will be found in the error term. If these cross-country differences are correlated with our explanatory variables (per capita GDP, per capita growth, Rule of Law), our OLS and IV coefficients will be distorted in the two equations.

16. They define a regime as a non-democracy in a given year if it meets one or more of the following four conditions: *i)* the chief executive is not elected; *ii)* the legislature is not elected; *iii)* there is no more than one party (this condition applies if there are no parties, if there is only one party, if the current term of office ends in the establishment of non-party or one-party rule, or if the incumbents

ISBN 92-64-02685-1 © OECD 2006

unconstitutionally close the legislature and rewrite the rules in their favour), or *iv)* none of the preceding conditions applies but the incumbents will have or already have held office continuously by virtue of elections for more than two terms (e.g. Singapore, Botswana, Mexico until 2000) or have held office without being elected for any duration of their current tenure in office, and until today or until the time when they were overthrown they had not lost an election.

17. Dollar figures are 1985 international prices (from Heston and Summers, 1993; Penn World Tables 5.6).

ISBN 92-64-02685-1 © OECD 2006

Chapter 6

Moving Forward

Summary

Building on findings of previous chapters — including that most governance indicators lack transparency and are not well suited to help developing countries identify how to improve the quality of local governance — this chapter briefly looks at promising new developments in the "market" for governance 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 both external users and developing countries seeking to improve the quality of local governance.

The remarkable explosion of growth in the demand for, and use of, governance indicators has been — and continues to be — driven largely by international investors and providers of official development assistance whose awareness of the importance of the quality of governance in developing countries for the success of their investments in or assistance to those countries has greatly increased in recent years. Reflecting both the management maxim that "what you cannot measure you cannot manage" and the need for practical instruments to use in their business and aid-budget-allocation decision-making processes in the face of the complex and often poorly understood realities of governance in developing countries, both groups rely primarily on *perceptions*-based *composite* governance indicators.

ISBN 92-64-02685-1 © OECD 2006

Perceptions-based Indicators

Two reasons go far to explain the importance of perceptions-based indicators, relative to that of facts-based indicators: *a)* the data required to construct facts-based indicators are often lacking, or lacking credibility, for developing countries; and *b)* the gap between the *de jure* "facts" which facts-based indicators often use and the informal, largely unwritten, institutions which they tend to ignore but often weigh much more heavily in *de facto* governance realities in developing countries, means that facts-based indicators can be seriously misleading. Some early debate focussed on the issue of whether users should rely more on facts-based indicators, despite these problems, because of the inherently subjective and non-replicable nature of perceptions-based indicators. And indeed, from a strictly scientific perspective, the fact that users can normally replicate facts-based but not perceptions-based indicators constitutes an important advantage of the former over the latter. But recognition of the significant elements of subjectivity inherent in the selection and interpretation of data used in the construction of facts-based indicators, added to the problems of data shortages and the gap between formal and informal institutions, has largely resolved this debate in favour of a consensus that the two types of indicators should be understood not as substitutes but as potentially useful complements that supply different types of information.

Moreover, from a practical investment decision-making perspective, the fact that many perceptions-based indicators reflect mainly business-oriented perceptions, from both inside and outside the country, can be of considerable value in its own right, i.e. for investors, whether or not those perceptions correctly reflect the quality of local governance as seen from the perspective of non-business interests (e.g. perspectives that are more adequately reflected in household surveys). The fact that users tend to rely on the same indicators which they see their peers using (and this is as true of other users as it is of business users) has led to a veritable bubble effect. The risk, of course, as implied by the US business spokesperson (who noted that a serious shortcoming of governance indicators for businesspeople is that in using them "we are talking mainly to ourselves"), and one that is analogous to the problem of country-risk rating systems explained in Chapter 3, is that perceptions-based governance indicators can be subject to herd behaviour by the "experts" on whose assessments they tend to rely, as noted for the KKZ indicators in Chapter 4. These indicators may therefore provide too little objective information about the quality of local governance for the very needs of local and foreign business decision makers.

ISBN 92-64-02685-1 © OECD 2006

Indeed, as others have also noted, perceptions-based indicators typically reflect perceptions of governance *outcomes*, not their causes or mechanisms. A "rule of law" indicator, for example, tells us how secure business people feel about their property, but little about precisely what makes them feel that way. Thus, as a former World Bank country director noted about the Bank's governance prescriptions, which rely on such governance indicators, they "come very close to a tautology. What is required for growth? Good governance. And what counts as good governance? That which promotes growth" (B. Kavalsky cited in *The Economist,* 2005).

Composite Indicators

The widespread use of *composite* indicators in turn reflects an understandable tendency on the part of users to reduce the complex realities of governance to a single number, for a given country in a given year, in order to facilitate a comparison of the quality of governance among countries and/or over time. Yet the very techniques used to quantify and synthesize a diversity of qualitative governance features into a single number make it difficult meaningfully to compare those numbers either across countries or over time — with the partial exception, illustrated by the "rule of thumb" developed by Kaufmann *et al.* as explained in Chapter 4, of using indicators only to compare scores across countries and over time that are *significantly* different from one another. Our analysis of how governance indicators tend widely to be used points to a great deal of misuse in this regard, by academics and other development analysts as much as by international investors and providers of development assistance.

Transparency Paradox

Particularly serious from the perspective of using governance indicators as a policy tool to help *improve* the quality of local governance is the fact that composite indicators tend to be characterised by a lack of transparency. One reason for this lack of transparency is that most of the widely used composite indicators are perceptions-based indicators and, as such, are both not replicable by users, and determined by subjective views of which users cannot be fully aware. A second reason is of course the complexity of the information that a

ISBN 92-64-02685-1 © OECD 2006

composite indicator conveys with a single number — complexity due to the diversity and often large number of underlying indicators used in its construction — which is the very *raison d'être* of composite indicators. A third reason is the absence of a clear underlying conceptual framework and a lack of clarity of the precise criteria for scoring. A fourth is the difficulty for users to gain access to some of the underlying indicators used to construct a given composite indicator — whether because they are expensive or simply not available for public usage.

The importance many aid donors attach to using governance indicators precisely to give *transparency* as well as coherence to their aid-budget-allocation decisions, in which they pay growing attention to the quality of governance in potential recipient countries, means that this lack of transparency of the governance indicators constitutes a serious problem. It is paradoxical, to say the least, for donors and investors to judge and sometimes punish developing countries for a perceived lack of transparent governance on the basis of such non-transparent indicators.

Internal and External Stakeholders

Also important to underline is the fact that from a developing-country perspective, the active users of governance indicators for decision-making purposes are overwhelmingly *foreign* investors and donors — referred to in the literature as "external stakeholders". Already by the late 1990s, concern was emerging that developing countries' own governments, business associations, NGOs and other such "internal stakeholders" were largely unable to use governance indicators to help bring about actual *improvement* in the quality of governance in their countries.

The reasons are largely the same: The lack of transparency, lack of underlying conceptual framework or theory of governance to help identify the *causes* of the governance outcomes reflected in the indicators, lack of clarity of the precise reasons why a country receives a particular score on a given indicator — all these features of the most widely used governance indicators undermine their usefulness as potential policy tools for developing countries. To bring about needed improvements in governance, internal stakeholders need indicators that can help them identify specific governance problems and monitor progress in reform with a degree of precision that is well beyond that of most governance indicators currently in use.

ISBN 92-64-02685-1 © OECD 2006

Moreover, while greater transparency is necessary for most indicators to be useful as policy tools, it is not sufficient. A good illustration is the set of facts-based governance indicators used by Przeworski, *et al.* (2000) cited in Chapter 5. These indicators are quite transparent, but they are not likely to be useful for identifying specific governance problems to be overcome in a given country or for closely monitoring the success of efforts to do so. The reason, of course, is that the indicators serve to define discrete political regime-types (whether a given country qualifies in a given year as a "democracy", and if not whether it is an "autocracy" or a "bureaucracy"), categories which constitute an objective, replicable, basis for identifying associations between different regime-types and economic characteristics (average level of economic growth, demographic growth, capital intensity, etc.) but which do not reveal, any more than do perceptions-based indicators, the process or steps countries need to follow to overcome barriers to transforming themselves from one regime-type to another[1].

New Initiatives

Growing concern that governance indicators provide little guidance for efforts actually to improve the quality of governance in developing countries, together with a growing focus on the needs of internal stakeholders (and concern about too little local "ownership" of efforts to improve governance), have generated important new initiatives since the late 1990s. Like any other effort to measure governance, these new indicators of governance will be subject to non-trivial measurement errors and policy makers should be cautious not to over-interpret small differences in scores across countries and over time. The real value-added in these initiatives is therefore not that they are less subject to measurement error, but that they produce indicators which are both more transparent and concrete enough to be more directly useful for policy makers in developing countries to identify specific kinds of change that are needed. We briefly present four such initiatives, and include others in Box 6.1.

WBI's Governance Diagnostic Surveys

Daniel Kaufmann and his team at the World Bank Institute have developed country-specific "Governance Diagnostic Surveys" that involve separate surveys in a given country of: *a)* households as users of public services;

ISBN 92-64-02685-1 © OECD 2006

b) business managers; and *c)* public officials[2]. An important feature of these Diagnostic Surveys is that they allow for a triangulation of results to compare the perceptions of the three groups. Another is that many of the questions seek explicitly to draw on respondents' direct experience, rather than simply on their generic impressions[3]. Both the ability to differentiate among different population groups (including in different geographic regions) and the specificity and experienced-based nature of the questions make the results more helpful for governments to design targeted strategies to combat corruption and other governance problems. Thus, for example, results for Sierra Leone indicated that a much higher percentage of households (60 per cent) were paying bribes in trying to obtain a public service than were managers (20 per cent) in trying to obtain a license or permit; results for Peru indicated a much higher rate of bribery being paid by households trying to obtain a public service in the Andean *sierra* region (about 15 per cent) than by households in the tropical *selva* region (about 7 per cent); and those for Albania showed public officials reporting that more than 50 per cent of customs inspectors "purchase" their positions, which helped the government design an Action Plan in 2003 to reform staff recruitment, improve the auditing of customs warehouses and enhance personnel policy by systematically carrying out individual staff evaluations and rotating staff every two to three years.

Results such as those cited here from Sierra Leone and Peru must of course be *interpreted* with care, not least because different respondent groups may have different fears or other motivations to under- or over-state their true experience, for example. Our point is simply that these country-specific Diagnostic Surveys, launched in the late 1990s, are very useful for policy makers and others seeking to *improve* the quality of governance in a given country — probably more so than are the KKZ and other composite indicators that are now widely used, and misused, by development analysts and academics as much as by international investors and official development agencies.

TI's Global Corruption Barometer

In 2003, Transparency International launched an annual household survey, covering up to 50 000 people in 64 countries, which is carried out for TI by Gallup International. Two important features of this public opinion survey, called the "Global Corruption Barometer", are that TI provides the disaggregated data (per question and per country) free of charge on its webpage, and that the results are comparable across countries and over time. These results include information on how people report their direct experience

ISBN 92-64-02685-1 © OECD 2006

of corruption, as well as their perceptions of corruption and expectations of future corruption. The data distinguish between corruption in public and private institutions and between petty and grand corruption, which allows users to assess where households report corruption's impact to be greatest[4].

DIAL (Development, Institutions and Long-term Analysis)

This public research organisation based in Paris undertakes household surveys to assess governance and democracy in Sub-Saharan Africa and Latin America in co-operation with national statistical offices in the countries where they work. These surveys include many experience-based questions[5] designed to produce results that are comparable both across countries and over time. DIAL undertakes "mirror" surveys in which "experts" and local households are asked the same questions and their answers compared. Furthermore, DIAL asks the experts what they believe the answers of the local households to be. Among their findings, for example from a cross-country comparison of perceptions of corruption in Africa, are that while the experts expected on average that 32 per cent of the local population would believe bribery to be an acceptable practice, only 5 per cent of the local population actually reported holding this belief; and while the experts believed that 54 per cent of the population would say it had experienced acts of corruption during the preceding year, only 13 per cent reported it had. Since the experts' assessments reported in these mirror surveys are largely similar to those embodied in the widely used perceptions-based indicators we have discussed throughout this study, DIAL's results raise further serious questions about the degree of inaccuracy and bias likely to be embodied in those indicators.

Metagora

DIAL is also part of the so-called "Montreux-Munich-Brussels Process" (InWEnt, 2004) which brings together users and producers of quantitative indicators of governance, democracy and human rights. The process was launched in 2000 with a conference in Montreux, Switzerland, on "Statistics, Development and Human Rights". Held under the auspices of the International Association for Official Statistics[6] and organised by the Swiss Federal Statistical Office and the Swiss Development and Co-operation Agency with the support of 15 organisations[7], this conference was followed in 2002 by three further meetings, two organised by the European Commission (Eurostat) together

ISBN 92-64-02685-1 © OECD 2006

with InWEnt[8]: *a)* in Munich on "Measuring Democracy and Good Governance", and *b)* in Brussels on "Statistics and Human Rights", and one organised by the Mexican National Commission for Human Rights with the support of Switzerland: the "International Seminar on Indicators and Diagnosis on Human Rights" in Merida. These were followed in 2005 by *c)* the New Delhi meeting on "Engendering and Empowering Governance Indicators" organised by the UNDP Oslo Governance Centre together with the Indian Council on Social Science Research and *d)* the Paris meeting on "Measuring Democracy, Human Rights and Governance" organised by Metagora[9].

Metagora is a pilot project, designed in the follow-up of the Montreux Conference and supported by the European Commission together with the governments of France, Sweden and Switzerland. It is implemented under the auspices of PARIS21, an OECD-hosted consortium whose aim is to foster more effective dialogue between producers and users of statistics on development issues. Metagora works at the national level in developing countries with a "bottom up" approach, involving international partners (e.g. DIAL) along with local partners, including local statistics authorities. Its purpose is to help diverse local stakeholders develop methods and tools with which to produce the data and indicators they need to formulate or evaluate specific national policies to promote democracy, human rights and governance. It is particularly active in francophone Africa, South Africa, the Andean Community, Mexico, Sri Lanka and the Philippines, and with Palestinians.

Box 6.1. **More New Initiatives to Produce Governance Indicators**

UNDP: The UNDP Oslo Governance Centre plans to assist 6 to 8 pilot countries in producing non-ranking "core" governance indicators which reflect universal aspects of governance relevant in all countries at both the national and local levels, and more country-specific "satellite" indicators, all to be used for pro-poor and gender sensitive policy reform. See http://www.undp.org/oslocentre/cross.htm.

UNU and ODI: The United Nations University undertook the first phase of a "World Governance Assessment" between 2002 and 2002 and a second phase by the United Kingdom's Overseas Development Institute is planned for 2006. The methodology and complete survey results of the first phase, based on responses from some 100 parliamentarians, civil servants, government officials, business people, academics and NGOs in 16 developing and transition countries, are publicly available. See http://www.odi.org.uk/wga_governance/About_WGA.html.

ISBN 92-64-02685-1 © OECD 2006

Box 6.1 (contd.)

MINEFI: France's Ministry of the Economy, Finance and Industry has compiled an Institutional Profile Database from surveys of staff at official French economic missions in 51 countries, and in 2006 is undertaking an identical survey, plus additional questions, in the missions in 85 countries. The questionnaire and the disaggregated data are available free of charge. See http://www.cepii.fr/francgraph/bdd/institutions.htm.

OECD: The OECD Directorate for Public Governance and Territorial Development plans in the coming years to develop and compile transparent, mainly facts-based, indicators of efficient public services in OECD countries to serve as a basis for diagnosing specific civil-service problems in countries, for monitoring the success of reforms, and for comparing countries. The initial focus will be on inputs and processes (e.g. budget processes, nature of the civil, service, structure of government, etc.), with data on the nature and scope of public services and their quality (e.g. processing time, customer/citizen focus, accessibility) to be gathered in a second phase. See http://www.oecd.org/gov/indicators.

Transparency of Indicators: An Imperative

An important challenge for new governance indicators is to gain acceptance and become widely used in a field that is overcrowded with existing indicators. Their success will depend on their use by decision makers. While most users rely on indicators they see the majority of their peers using — there has been a real bubble effect as we noted earlier — our recommendation to users is not to join the herd. Assess alternative sets of indicators critically to determine their true suitability for your needs.

The criteria for determining the usefulness of different governance indicators can be expected, of course, to vary according to their user's purpose. One can also only agree with those like the UNDP, in their *Governance Indicators: A Users' Guide* (Sudders and Nahem, 2004), who strongly urge users to "use a *range* of indicators" — which it describes as users' "Golden Rule One" — on the grounds that, "A single governance indicator which captures the subtleties and intricacies of national situations…does not exist. Using just one indicator could very easily produce perverse assessments of any country and will rarely reflect the full situation." Yet it is equally true, as the UNDP also immediately notes, that "having too many indicators results in a different range of problems,

ISBN 92-64-02685-1 © OECD 2006

including a lack of focus and burdensome data collection and analysis". What is their proposed solution? "The key," they say, is for users to employ "a balanced set [of indicators] with sufficient but not superfluous information."

We believe it even more important to emphasise the need for greater *transparency* in the production and use of governance indicators. Specifically, we believe it possible, and important, for users, especially "external stakeholders", to raise the minimum-quality standard they set for the governance indicators they use — particularly as regards the transparency of those indicators — relative to those widely in use today. Our reasoning is as follows.

— Recent and growing efforts to respond more effectively to the needs of "internal stakeholders" are now generating in-depth and relatively transparent governance-related information, and indicators, which can also be very useful for international investors, providers of ODA, academics and other development analysts.

— Answers to more experience-based questions, together with triangulated and "mirror" surveys that objectively differentiate the assessments of "experts" from those of other important and clearly defined population groups, can be used to reduce the likely inaccuracies and biases embodied in composite perceptions-based indicators and provide other valuable information for users — internal and external alike — who wish to make comparisons across countries and/or over time.

Our recommendation is therefore to move beyond the distinction between relatively transparent indicators built to serve "internal stakeholders" and the policy needs of developing countries, on the one hand, and composite perceptions-based indicators widely used — and misused — by "external stakeholders" for cross-country comparisons, on the other. Transparent and publicly available sets of governance indicators based on facts and/or the perceptions of a *diversity* of clearly defined population groups both within and outside countries can be aggregated for broad cross-country comparisons of greater use, and accuracy, for all groups[10]. They should be aggregated with a technique that does not effectively attribute negligible weights to sources that differ from the majority (see also Box 4.2)[11]. Not only would such transparent and publicly available indicators be useful for the purposes of diverse user groups, they are likely to foster fruitful discussion and greater understanding of the *causal mechanisms* that lie behind governance problems both within and across countries.

To achieve such greater understanding of the causes of governance realities requires users effectively to *understand* the strengths and limitations of the indicators they use — as the UNDP also emphasises in making such

ISBN 92-64-02685-1 © OECD 2006

understanding their "Golden Rule Three". This understanding in turn requires, in our view, achieving a degree of transparency in the production of indicators that fulfils two criteria:

— *Full disclosure of methods and criteria for scoring:* Producers should clearly explain for non-specialists the methodology used to construct their indicators, including underlying assumptions that give *de facto* "meaning" to specific indicators (e.g. stricter labour or environmental regulations are assumed to lower the quality of governance) along with a full list of underlying indicators incorporated in the construction of composite indicators, and clarification of the purposes for which data used to produce each underlying indicator were originally collected.

— *Full disclosure of countries' scores:* Producers should publish not only countries' scores on each composite indicator but on each of the underlying indicators from which the composite indicator is constructed.

Meeting this standard of transparency in the production and use of governance indicators is not only important to help overcome the damaging reality, resented today by many developing countries, that they are pressured to raise their standards of transparency in governance by OECD-based suppliers of funds who use highly un-transparent indicators to judge them. Much more positively, and ultimately more importantly we believe, is the likelihood that raising the level of transparency embodied in the governance indicators that international investors, providers of aid, analysts and academics, as well as domestic policy makers and other "internal" stakeholders all use, will greatly facilitate discussion among them on how best to bring about needed *improvement* in the quality of local governance. In doing so it will also probably reveal how *little* we still understand about the institutions needed for good governance in developing countries — reflected as well in the lack of a recognised theory of governance from which to construct governance indicators.

Encouraging organisations to clarify the norms and criteria behind their standards for "good governance" by setting higher standards for the transparency of the indicators they use should thus help to strengthen the bases, both analytical and political, for better understanding the kinds of institutions most conducive to development in developing countries today. This should in turn help raise awareness both locally and internationally of how best to try to overcome obstacles to achieving those institutions. Raising the standards of transparency that we demand in the governance indicators we use, and produce, should thus help effectively to improve the quality of governance in developing countries.

ISBN 92-64-02685-1 © OECD 2006

Notes

1. Another prominent example of a more transparently constructed set of governance indicators is the Polity Project that was put into place in 1975 by Ted Gurr at the University of Maryland. Ted Gurr and his associates have assembled a large historical data set (Polity I, II, and III, IV) which in its latest version comprises 161 contemporary countries with populations of more than 500 000 inhabitants reaching back in the oldest states to about 1800, or to the year of independence for the more recent ones (M.G. Marshall, K. Jaggers and T. R. Gurr, 2005). Based on a variety of sources for each country, including objective data such as the countries' respective constitutions, the authors use a particularly transparent coding method to produce indicators on the competitiveness and openness of executive recruitment, the constraints on the chief executive, the regulation of participation and the competitiveness of political participation, and aggregate these "authority variables" into a democracy indicator, an authority indicator, and a combined Polity Indicator. This information on the authority characteristics of most countries in the world, on an annual basis, is very useful for both cross-country studies and studies of change over time. But it is not obvious how policy makers can use these indicators to identify the specific kinds of change required actually to improve governance in their country.

2. As of early 2006, WBI Governance Diagnostic Surveys had been carried out in Albania, Benin, Bolivia, Brazil, Cambodia, Ecuador, Ghana, Guatemala, Guinea, Haiti, Honduras, Indonesia, Latvia, Madagascar, Mozambique, Paraguay, Peru, Sierra Leone, Slovakia and Thailand (see http://worldbank.org/wbi/governance/capacitybuild/diagnostics.html).

3. A generic question, of the type often addressed to "experts" whose assessments supply the data used in KKZ and other perceptions-based indicators, would be, "Rate the level of corruption in your country on a scale of 1 to 7", whereas typical experienced-based questions are, "In 2005, on average, what percentage of the time of the director and managers of a business like yours was spent negotiating with public officials?" and, "In 2005, on average, what do you believe to be the total percentage of profits that a business like yours spends dealing with public officials in public and private affairs, meals, parties, etc.?"

4. See http://www.transparency.org/policy_research/surveys_indices/gcb

ISBN 92-64-02685-1 © OECD 2006

5. Using as proxies for actual corruption households' perceptions on corruption that do not relate directly to households' concrete personal experiences can produce misleading results. Olken (2006) compares a measure of missing expenditure in the context of a road-building programme in rural Indonesia with the perceptions of the villagers on the likelihood of the diversion of money in the road project. He finds that ethnically heterogeneous villages have higher perceived corruption levels but lower levels of missing expenditures. He suggests that this reverse correlation may be explained by the lower levels of trust and greater citizen monitoring observed in heterogeneous villages.

6. A chapter of the International Statistical Institute (ISI), the International Association for Official Statistics is an NGO bringing together producers and users of official statistics that presents itself as "an open and universal forum where to regularly consider the development of official statistics and in particular its role and contribution in a democratic society" (http://www.stats.govt.nz/iaos/home.htm).

7. These include four UN bodies (UNICEF, UNECA, UNHCR, UNFPA), the International Labour Office, the European Commission (Eurostat), the Council of Europe, the European Free Trade Association, university-based centres and others.

8. InWEnt is *Internationale Weiterbildung und Entwicklung* (International Capacity Building), a joint undertaking by Germany's Federal Government, state (Länder) governments and industry to provide education, exchange and dialogue programmes every year for about 55 000 specialists, executives and other decision makers in industry, politics, administration and civil society from developing, transition and OECD countries.

9. Further important meetings include the "The Conference on Measuring Governance" sponsored by the World Peace Foundation and the Program on Intrastate Conflict, Conflict Prevention, and Conflict Resolution at Harvard University's John F. Kennedy School of Government in May 2003, and two workshops on the Measurement of Human Rights organised by the Carr Center for Human Rights Policy at the Kennedy School with the support of the Macarthur Foundation in May 2005 and July 2006.

10. According to the World Bank's Global Monitoring Report (2006), some progress is being made toward developing composite indicators on the basis of disaggregated "action-oriented" indicators.

11. The OECD's Statistic Directorate (Nardo *et al.*, 2005) has published a "Handbook on constructing composite indicators: methodology and user guide". The handbook aims to provide a guide for constructing and using composite indicators for policy makers, academics, the media and other interested parties. It discusses different aggregation methods and shows through examples that the choice of aggregation method heavily influences countries' positions on indicators. While there is no perfect aggregation method, the handbook suggests to perform a sensitivity analysis to find out how changes in the aggregation method and in the inclusion and exclusion of sub-indicators affect countries' scores and their relative position with respect to other countries.

ISBN 92-64-02685-1 © OECD 2006

Appendix I
Aggregation Methodology for the KKZ Composite Indicators

We first describe the model and the assumptions underlying the aggregation procedure, then we explain the aggregation procedure step by step.

The Model

Kaufmann, Kraay and Mastruzzi use an unobserved components model:

1. $y(j,k) = \alpha(k) + \beta(k)(g(j) + \varepsilon(j,k))$

List of Abbreviations and Indices

j	Country, j = 1,2….J
k	Indicator (subcomponent of the composite indicator), k=1,2…K
y (j,k)	observed score on indicator k for country j
g(j)	Unobserved "true governance", in our example true Rule of Law. g(j) is postulated to exist in the form of a normally distributed random variable with mean 0 and standard deviation 1.
ε(j,k)	Disturbance term also referred to as error term. It consists of not only perception and measurement error and sampling variation, but also the imperfect relationship between the particular concept measured by indicator k and the corresponding broader aspect of governance. Judicial independence and crime as measured by World Markets Online for example might be imperfect proxies for Rule of Law.
α(k), β(k)	Coefficients serving to map, together with the disturbance term ε(j,k), unobserved governance into the observed data.
$\sigma_{\varepsilon}^{2}(k)$	Variance of the disturbance terms of indicator k common to all countries j

ISBN 92-64-02685-1 © OECD 2006

Thus, if j = Singapore and k= the indicator from World Markets Online for Rule of Law, then y(j,k) is the observed score on the World Markets Online indicator for Rule of Law in Singapore and g(j) is the "true" Rule of Law in Singapore, the variable we are interested in.

The unobserved components model is used to account for the fact that governance itself is not observable and that we can only approximate it by aggregating the scores we obtain on each indicator, y(j,k). In order to be able to estimate the unknown "true governance", g(j), it is convenient to put the error term together with g(j) into brackets.

The Assumptions Underlying the Model

1) The random terms $\varepsilon(j,k)$, called the disturbance terms, are uncorrelated with each other, i.e. perception errors are uncorrelated across indicators and countries. Some assumption of this type is necessary for the identification of the model parameters. The mean of $\varepsilon(j,k)$ is zero for all j,k.

2) The disturbance term has the same variance, $\sigma_{\varepsilon}^{2}(k)$, across countries within a given indicator but may have a different variance across indicators.

3) The relationship between unobserved governance and observed indicators is linear.

4) The disturbance terms $\varepsilon(j,k)$ are statistically independent of the unobserved components g(j) for all j and k.

5) Unobserved governance g(j) have a joint normal distribution and the disturbance terms $\varepsilon(j,k)$ have a joint normal distribution.

Estimating the "True" Level of Governance

The "true" level of Rule of Law in a specific country is unknown. We can only estimate it conditionally on the results we obtain for each indicator in the Rule of Law Cluster. These indicators are aggregated in order to estimate the "true" level of Rule of Law.

ISBN 92-64-02685-1 © OECD 2006

First of all, Maximum Likelihood estimates of $\alpha(k)$, $\beta(k)$ and $\sigma_\varepsilon(k)$ are obtained.

This is based on the following Likelihood function:

2. $$L\left(y;\alpha,\beta,{\sigma_\varepsilon}^2(1),...,{\sigma_\varepsilon}^2(K)\right)=\prod_{j=1}^{J}(2\cdot\pi)^{-J/2}\left|\Omega\right|^{-1/2}\cdot\exp(-\frac{1}{2}\cdot\left(y(j)-\alpha\right)'\Omega^{-1}\left(y(j)-\alpha\right))$$

K = number of indicators

J = number of countries

y(j) = the Kx1 vector of the y(j,k)'s for country j

y = the JKx1 vector of the y(j,k)'s for all countries

α = Kx1 vector of the α(k)'s

β = Kx1 vector of the β(k)'s

Ω = $\beta\beta'$ + diag $\{\sigma_\varepsilon^2(k)\cdot\beta(k)^2\}$

The weights for each indicator in the aggregation procedure are inversely proportional to its error variance, i.e. the greater the variance of the error term the smaller the weight. They are given by:

3. $$w(k)=\frac{\sigma_\varepsilon(k)^{-2}}{1+\sum_{k=1}^{K(j)}\sigma_\varepsilon(k)^{-2}}$$

Kaufmann *et al*. base their estimate of "true" governance on a weighted average of the rescaled observed scores with the weights w(k) functions of relative error variances (equation 3). Rescaled means that α(k) is subtracted from each observed score y(j,k) and the result is then divided by β(k). This is based on rewriting formula 1 and taking a mathematical expectation. We are dealing with a calculation of expected values and the expected value of the disturbance term, ε(j,k), is by assumption 0.

The conditional distribution of unobserved governance g(j) is normal as a consequence of the assumptions with the following mean (Equation 4) and the following standard deviation (Equation 5).

4. $$\mathrm{E}\left[\mathrm{g(j)}\middle|\mathrm{y(j)},\alpha,\beta\right]=\sum_{k=1}^{K(j)}w(k)\cdot\frac{y(j,k)-\alpha(k)}{\beta(k)}$$

ISBN 92-64-02685-1 © OECD 2006

5. $$sd\left[g(j)\middle|y(j),\sigma_{\varepsilon}{}^{2}(1),\ldots,\sigma_{\varepsilon}{}^{2}(K)\right]=\left[1+\sum_{k=1}^{K(j)}\sigma_{\varepsilon}(k)^{-2}\right]^{-1/2}$$

This standard deviation declines in the number of individual indicators in which a particular country appears and increases in the variance of the disturbance term on each of these indicators.

The following section describes the estimation of the "true" level of governance (Equation 4) and its standard deviation (Equation 5) step by step.

The Aggregation Procedure Step by Step

The several hundred indicators from the 37 data sources are assigned to the 6 governance categories: "Voice and Accountability", "Political Stability" and "Absence of Violence", "Government Effectiveness", "Regulatory Quality", "Rule of Law" and "Control of Corruption".

Each of the 6 aggregate indicators is then constructed in 11 steps. For the purpose of illustration, the calculation of the Rule of Law indicator is described:

1) All indicators from the same source in the Rule of Law cluster are combined into a single indicator. The source World Markets Online for example provides two indicators that are relevant for "Rule of Law":

Judicial Independence —"*an assessment of how far the state and other outside actors can influence and distort the legal system. This will determine the level of legal impartiality investors can expect.*" — and

Crime — "*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.*" (Kaufmann *et al.*, 2003, p.87).

A simple average of these indicators is taken. As we have combined all indicators from the same source into one indicator, we have now as many indicators as sources in the Rule of Law cluster, namely 24.

2) Each indicator constructed in this way is rescaled, so that higher outcomes correspond to better outcomes. A further rescaling is achieved by first subtracting the minimum *possible* score and then dividing by the difference between the minimum and maximum *possible* score.

ISBN 92-64-02685-1 © OECD 2006

3) Each of the 24 sources for "Rule of Law" is classified as representative or non-representative depending on the distribution of the included countries across geographical regions and income. For "Rule of Law", nine of the ten representative sources are expert assessments.

4) For the representative indicators in the Rule of Law cluster, the estimates of the unknown parameters, $\alpha(k)$, $\beta(k)$ and $\sigma_{\varepsilon}^{2}(k)$, are obtained using Maximum Likelihood Estimation (Equation 2)[1]. The indicators that obtain similar result to the other indicators will have a low estimated $\sigma_{\varepsilon}^{2}(k)$, whereas indicators that are not correlated with the other indicators will have larger error variances. The reasoning behind this calculation is that the correlation of scores is due to the same underlying concept of "true" governance and not due to a correlation of perception errors, an assumption that is discussed below.

5) Weight calculation: Thus, highly correlated indicators will have a larger weight than other indicators, as the weights assigned to each indicator (equation 3) are inversely proportional to their imputed error variance.

6) Equation 4 can now be estimated for each country using only the representative sources, so that an estimation of the "true" level of Rule of Law, g(j), is obtained for each country.

7) The standard error of these estimates is calculated (Equation 5).

8) The non-representative indicators are regressed on these estimates of unobserved governance to obtain estimates of $\alpha(k)$, $\beta(k)$ and $\sigma_{\varepsilon}^{2}(k)$, for the non-representative indicators. They are corrected for the attenuation bias imparted by measurement error in the estimates of unobserved governance obtained in Step 6 by using the standard error obtained in Step 7.

9) Drawing on all indicators, representative and non-representative, new weights (Equation 3) and based on that, new estimates for the "true" level of Rule of Law for each country (Equation 4), can be calculated

10) The estimates of "Rule of Law" for each country are rescaled by subtracting the mean across countries and dividing by the standard deviation across countries. Almost all scores are now in the range between -2.5 and 2.5.

11) The standard error (Equation 5) is recalculated.

ISBN 92-64-02685-1 © OECD 2006

What consequences does the violation of the assumption that disturbances are not correlated have?

The assumption allows us to identify the portion of the variation in scores across countries within each indicator caused by measurement error. An indicator that is highly correlated with the other indicators is interpreted as having a small residual variance (Kaufmann *et al.*, 2003). If errors are correlated, the covariance terms in the maximum likelihood function are improperly set to zero. This invalid constraint imposed on the likelihood function makes inconsistency likely. The Maximum Likelihood Estimation procedure is likely to lead in this case to inconsistent estimates of alpha, beta and the error variance[2]. The estimates for governance obtained in step 1 to 6 would therefore be inconsistent and the standard error would be higher. Step 8, the regression of the non-representative indicators on these estimates for governance would therefore produce inconsistent estimates for alpha, beta and the error variance and can also not be appropriately corrected for measurement error in the estimates for governance, as the standard error obtained in Step 7 is probably underestimated.

In sum, in case of correlated errors, we suspect that the aggregation procedure will produce biased and inconsistent estimates. Kaufmann *et al.* warn that the reported margins of errors are already quite high assuming the best-case scenario, i.e. the non-correlation of the disturbance term. In the likely case of correlated disturbance terms across indicators, the margins of errors will be even higher, making the cross-country classification even more difficult.

Notes

1. The reason why MLE cannot be applied to the non-representative sources is that the distribution of unobserved governance in the subset of countries covered by these sources is different than the distribution in the whole set of countries. g(j) will therefore not follow a standard normal distribution which is required for (this application of) MLE.

2. The likelihood function specified contains a diagonal matrix assuming that the covariance of the error variances equals 0 which under Normality implies independence of disturbances. It would be interesting to use the General Method of Moments approach to allow for correlation of disturbance terms and therefore a non-diagonal covariance matrix (as well as deviations from Normality).

ISBN 92-64-02685-1 © OECD 2006

Appendix II
Governance and Growth

List of Variables

Variable	Description and Source
Africa	Regional Dummy for Africa
Cars78to81	Passenger cars (per 1 000 people), data either from the year 1978, 1979, 1980 or 1981, World Development Indicators, also available cars 1990 and cars1996
Cc	Control of Corruption 2000, Kaufmann *et al.* composite indicators, ranges from -3 to + 3.
East Asia	Regional Dummy for East Asia
Eastern Europe	Regional Dummy for Eastern Europe
EngFrac	the "first" language variable, corresponding to the fraction of the population speaking English, Hall and Jones (1998)
EurFrac	the "first" language variable, corresponding to the fraction of the population speaking one of the major languages of Western Europe: English, French, Germar Portuguese, or Spanish, Hall and Jones (1998)
Fitlnsetmor	Ln of Settler mortality, (Acemoglu, Johnson and Robinson, 2001)
Gdp	GDP Levels: PIB, (million 1990 International Geary-Khamis), Maddison (2003)
Gdpc	Per capita GDP: (1990 International Geary-Khamis dollars), Maddison (2003)
Ge	Government Effectiveness, Kaufmann *et al.* composite indicators, ranges from -3 to + 3.
Grc50s	Annual average growth rate in gdpc in the 50s
Grc70to2000	Annual average growth rate in gdpc from 1970 to 2000
Grpop70to2000	Annual average growth rate in pop from 1970 to 2000
Illit	Adult illiteracy rate, World Development Indicators
Infantmor	Mortality rate, infant (per 1 000 live births), World Development Indicators
Investrate	Average investment rate (gross capital formation) as percentage of GDP, World Development Indicators
Latin America	Regional Dummy for Latin America
Lgdpc	Ln of gdpc
Lives	Life expectancy at birth, total (in years). (WB 1994, series 387)
Lnyearsindp	Ln of years of independence, CIA, http://www.cia.gov/cia/publications/factbook/fields/2088.html
Pop	Population (000 at mid-year), Maddison
Rl	Rule of Law, Kaufmann *et al.* composite indicator. ranges from -3 to + 3.
Rq	Regulatory Quality, Kaufmann *et al.* composite indicator. ranges from -3 to + 3
West Asia	Regional Dummy for West Asia

Note: The years given in the table are only suggestive; we used Maddison's GDP per capita data for instance from 1900 to 2000.

ISBN 92-64-02685-1 © OECD 2006

Table A.II.1. **Explained Variable: Growth of Per Capita GDP**

Explanatory Variables	OLS (1)	IV (2)	IV (3)
Ln of GDP per capita 1970	-0.014	-0.029	-0.021
	(0.002)	(0.009)	(0.006)
Rule of Law 1996	0.011	0.040	0.026
	(0.002)	(0.017)	(0.010)
Average investment rate (gross capital formation) as % of GdP, wdi	0.128	0.106	0.115
	(0.019)	(0.033)	(0.026)
Growth rate in the population 1970-2000	-0.620	-0.376	-0.534
	(0.157)	(0.298)	(0.222)
Regional dummies (Western countries base level)			
Africa	-0.008	0.023	0.002
	(0.007)	(0.021)	(0.012)
East Asia	0.009	0.032	0.014
	(0.006)	(0.017)	(0.011)
Eastern Europe	-0.005	0.038	0.005
	(0.007)	(0.027)	(0.014)
Latin America	0.008	0.048	0.021
	(0.006)	(0.025)	(0.014)
West Asia	0.009	0.034	0.014
	(0.007)	(0.019)	(0.011)
Instruments			
Settler mortality		X	X
Adult illiteracy rate 1970			X
Ln of years of independence			X
Summary statistics			
Number of observations	128	127	107
R-Squared	0.639	0.007	0.467
SER (Standard Error of the Regression)	0.013	0.021	0.016

Notes:
(1) reg grc70to2000 lgdpc1970 grpop70to2000 investrate rl1996 Reg*[1].
(2) ivreg grc70to2000 lgdpc1970 grpop70to2000 investrate (rl1996=fitlnsetmor) Reg*.
(3) ivreg grc70to2000 lgdpc1970 grpop70to2000 investrate (rl1996=illit1970 fitlnsetmor lnyearsindp) Reg*[2].

ISBN 92-64-02685-1 © OECD 2006

Table AII.2. **Explained variable: Rule of Law 2002**

Explanatory Variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IV (5)	IV (6)	IV (7)	IV (8)
Ln of GDP per capita 1960 (million 1990 International Geary-Khamis dollars)	0.484 (0.097)				1.180 (0.275)			
Ln of GDP per capita 1970 (million 1990 International Geary-Khamis dollars)		0.538 (0.090)				1.110 (0.210)		
Ln of GDP per capita 1980 (million 1990 International Geary-Khamis dollars)			0.609 (0.084)				0.999 (0.155)	
Ln of GDP per capita 1990 (million 1990 International Geary-Khamis dollars)				0.738 (0.074)				0.963 (0.122)
Ln of Settler mortality	-0.232	-0.190	-0.134	-0.062	-0.052	-0.022	0.008	0.037
	(0.084)	(0.082)	(0.080)	(0.072)	(0.120)	(0.111)	(0.098)	(0.085)
Regional dummies								
(Western countries base level)								
Africa	-0.853	-0.750	-0.616	-0.302	0.113	0.077	-0.028	0.036
	(0.283)	(0.273)	(0.260)	(0.232)	(0.482)	(0.414)	(0.341)	(0.282)
East Asia	-0.681	-0.566	-0.517	-0.404	0.411	0.382	0.098	-0.074
	(0.269)	(0.259)	(0.240)	(0.207)	(0.520)	(0.437)	(0.333)	(0.259)
Eastern Europe	-1.218	-1.181	-1.156	-1.291	-0.668	-0.732	-0.869	-1.152
	(0.337)	(0.325)	(0.308)	(0.175)	(0.450)	(0.405)	(0.346)	(0.191)
Latin America	-1.231	-1.140	-1.116	-0.862	-0.850	-0.752	-0.880	-0.688
	(0.233)	(.227)	(0.214)	(0.195)	(0.322)	(0.301)	(0.252)	(0.222)
West Asia	-0.876	-0.868	-0.872	-0.494	-0.770	-0.731	-0.791	-0.353
	(0.247)	(0.238)	(0.226)	(0.207)	(0.305)	(0.281)	(0.246)	(0.222)
Instruments								
Infant mortality rate (per 1 000 live births) 1960					X			
Infant mortality rate 1970						X		
Infant mortality rate 1980							X	
Infant mortality rate 1990								X
Summary statistics								
Number of observations	134	134	134	154	127	129	130	151
R-Squared	0.652	0.676	0.707	0.744	0.515	0.569	0.661	0.728
Standard Error of regr.	0.648	0.626	0.595	0.535	0.777	0.730	0.645	0.554

Notes:

(1) reg rl2002 lgdpc1960 Reg* fitlnsetmor.
(2) reg rl2002 lgdpc1970 Reg* fitlnsetmor.
(3) reg rl2002 lgdpc1980 Reg* fitlnsetmor.
(4) reg rl2002 lgdpc1990 Reg* fitlnsetmor[3].
(5) ivreg rl2002 (lgdpc1960=infantmor1960) Reg* fitlnsetmor.
(6) ivreg rl2002 (lgdpc1970=infantmor1970) Reg* fitlnsetmor.
(7) ivreg rl2002 (lgdpc1980=infantmor1980) Reg* fitlnsetmor.
(8) ivreg rl2002 (lgdpc1990=infantmor1990) Reg* fitlnsetmor[4].

ISBN 92-64-02685-1 © OECD 2006

Table A.II.3. **Rule of Law 2002**

Explanatory variables	IV (1)	IV (2)	IV (3)	IV (4)
Ln of GDP per capita 1970 (million 1990	0.950	0.840	0.835	1.034
International Geary-Khamis dollars)	(0.191)	(0.140)	(0.158)	(0.134)
Ln of Settler mortality	-0.095	-0.144	-0.089	0.022
	(0.095)	(0.104)	(0.095)	(0.094)
Regional dummies				
(Western countries base level)				
Africa	0.018		0.035	0.314
	(0.402)		(0.331)	(0.303)
East Asia	0.269		0.079	0.127
	(0.391)		(0.277)	(0.258)
Eastern Europe			-0.965	-1.016
			(0.296)	(0.198)
Latin America	-0.841		-0.928	-0.313
	(0.268)		(0.232)	(0.254)
West Asia	-0.095		-0.645	-0.400
	(0.095)		(0.258)	(0.239)
Instruments				
Infant mortality rate 1970	X	X		
Infant mortality rate 1980			X	
Infant mortality rate 1990				X
Life expectancy at birth, total (in years) 1970	X	X		
Passenger cars (per 1 000 people) 1978-1981			X	
Passenger cars (per 1 000 people) 1990				X
Summary statistics				
Number of observations	96	96	73	122
R-Squared	0.712	0.588	0.789	0.743
Standard Error of regr.	0.602	0.703	0.505	0.547
Test of overidentifying restrictions: chi-square	2.957	0.806	0.329	3.660
(p-value)	(0.086)	(0.369)	(0.567)	(0.056)

Notes:

(1) ivreg rl2002 (lgdpc1970 = infantmor1970 lives1970) Reg* fitlnsetmor[5].

(2) ivreg rl2002 (lgdpc1970 = infantmor1970 lives1970) fitlnsetmor.

(3) ivreg rl2002 (lgdpc1980 = infantmor1980 cars78to81) Reg* fitlnsetmor.

(4) ivreg rl2002 (lgdpc 1990 = infantmor1990 cars1990) Reg* fitlnsetmor[6].

ISBN 92-64-02685-1 © OECD 2006

Notes

1. We tested the endogeneity of the Rule of Law variable in the growth equation and found that Rule of Law is endogenous in the growth equation (the residual of the reduced form regression of Rule of Law on all predetermined variables is significant when added to the OLS growth regression). We therefore decided to instrument Rule of Law in line with Kaufmann and Kraay.

2. Stata™ code: Reg stands for regression, ivreg for instrumental variable regression. The first variable is the explained variable and the following variables are the explanatory variables. (rl1996= fitlnsetmor) means that the variable "Rule of Law 1996" is instrumented by "settler mortality".

3. We tested the endogeneity of the GDP per capita Variable in the Rule of Law equation and found that GDP per capita is endogenous in the Rule of Law equation (The residual of the reduced form regression of GDP per capita on all predetermined variables is significant when added to the OLS growth regression). We therefore decided to instrument GDP per capita.

4. Stata™ code.

5. The choice of the reported IV regressions is based on the number of available observations.

6. Stata™ code.

ISBN 92-64-02685-1 © OECD 2006

Bibliography

Acemoglu, D., S. Johnson and J. Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review*, Vol. 1(5).

Aron, J. (2000), "Growth and Institutions: A Review of the Evidence", *World Bank Research Observer*, Vol. 15(1), World Bank, Washington, D.C.

Bellver, A. and D. Kaufmann (2005), "Transparenting Transparency: Initial Empirics and Policy Applications", *World Bank Policy Research Working Paper*, World Bank, Washington, D.C.

Berg-schlosser, D. (2002), "Indicators of Democratization and Good Governance as Measures of the Quality of Democracy – A Critical Appraisal", *Working Paper*, International Sociological Association Congress 2002, Brisbane.

Berthelemy, J.C. and A. Tichit (2004), "Bilateral Donors' Aid Allocation Decisions — A Three-Dimensional Panel Analysis", *International Review of Economics & Finance*, Elsevier, Vol. 13(3).

Berthelier, P., A. Desdoigts and J. Ould Aoudia (2003), "Institutional Profiles: Presentation and Analysis of an Original Database of the Institutional Characteristics of Developing, in Transition and Developed Countries", *Working Paper*, French Ministry of the Economy, Finance and Industry, Paris.

Besançon, M. (2003), "Good Governance Rankings: The Art of Measurement", *World Peace Foundation Reports,* WPF, Vol. 36, Cambridge.

Bouchet, M.H., E. Clark and B. Groslambert (2003), *Country Risk Assessment: A Guide to Global Investment Strategy,* Wiley Finance Series, John Wiley & Sons, West Sussex.

Business and Industry Advisory Committee to the OECD, *Investment – BIAC Position on Incentives*, 5 November 2002.

Burnside, C. and D. Dollar (1997), "Aid, Policies, and Growth", *World Bank Policy Research Working Paper*, No. 1777, World Bank, Washington, D.C.

Burnside, C. and D. Dollar (2004), "Aid, Policies, and Growth: Revisiting the Evidence", *World Bank Policy Research Working Paper*, No. 3251, World Bank, Washington, D.C.

ISBN 92-64-02685-1 © OECD 2006

CAISSE DES DÉPÔTS ET CONSIGNATIONS (December 1998), *Emerging Regions*, Vol. 5. Paris.

CHANG, H.J. (2002), *Kicking Away the Ladder: Development Strategy in Historical Perspective*, Anthem Press, London.

CORNIA, G. and V. POPOV (eds). (2001), *Transition and Institutions: The Experience of Gradual and Late Reformers*, UNU/WIDER, Oxford University Press, Oxford.

DAC/OECD (1995), *Participatory Development and Good Governance*, Development Co-operation Guidelines Series, OECD, Paris, available online at: www.oecd.org/dataoecd/27/13/31857685.pdf

DIAL (2005), *Global Corruption Report*, Paris.

DIAL (2006), *Dialogue: La Lettre d'Information de Dial*, No. 24, Paris, available at: http://www.dial.prd.fr/dial_publications/PDF/Dialogue/Dialogue24.pdf

DOLLAR, D. and V. LEVIN (2004), "The Increasing Selectivity of Foreign Aid (1984-2002)", *World Bank Policy Research Working Paper*, No. 3299, World Bank, Washington, D.C.

EASTERLY, W. (2002), *The Elusive Quest for Growth*, MIT Press, Cambridge, Massachusetts.

EVIATAR, D. (2003), "Do Aid Studies Govern Policies or Reflect Them?" *New York Times*, July 26, 2003, New York, NY.

Financial Times (2004), "Hazards of Charting Corruption", 20 October.

GALENSON, W. (1959), *Labor and Economic Development*, Wiley, New York, NY.

GALTUNG, F. (2005), "Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices" *in Measuring Corruption*, Ashgate, United Kingdom.

GELB, A., B. NGO and X. YE (2004), "Implementing Performance-Based Aid in Africa: The Country Policy and Institutional Assessment", *World Bank Africa Region Working Paper Series*, No. 77, World Bank, Washington, D.C.

GIRVAN, N. (2002), "Problems with UNDP Governance Indicators: The Greater Caribbean This Week", available online at: www.acs-aec.org/column/index45.htm

GLOBERMAN, S. and D. SHAPIRO (2002), "Global Foreign Direct Investment Flows: The Role of Governance Infrastructure", *World Development*, Vol. 30.

GLOBERMAN, S. and D. SHAPIRO (2003), "Governance Infrastructure and U.S. Foreign Direct Investment", *Journal of International Business Studies*, Vol. 34.

GRIFFITH-JONES, S. and R. GOTTSCHALK (2006), "Cost of Currency Crises and Benefits of International Financial Reform" *in* KAUL, I. (ed.), *The New Public Finance: Responding to Global Challenges*, Oxford University Press, New York, NY, available online at: www.ids.ac.uk/ids/global/Finance/pdfs/SGJCostofcrises2.pdf

HALL, R. and C. JONES (1998), "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Working Paper*, Stanford University, No. 98-007, available online at: ssrn.com/abstract=3595

 ISBN 92-64-02685-1 © OECD 2006

HAUSMAN, J.A. (1983), "Specification and Estimation of Simultaneous Equations Models" in GRILICHES, Z. AND M.D. INTRILIGATOR (ed.), *Handbook of Econometrics*, Vol. I, North-Holland 1983.

HERMAN, B. (2004), "How Well Do Measurements of An Enabling Domestic Environment For Development Stand Up?", UNCTAD, Geneva, available online at: www.g24.org/003gva04.pdf

HESTON, A. and SUMMERS, R. (1993), *The Penn World Table 5.6.*, available on diskette from the National bureau of Economic Research, Cambridge, MA.

HUNTINGTON, S. and J. NELSON (1976), *No Easy Choice: Political Participation in Developing Countries*, Harvard University Press, Cambridge, Massachusetts.

INTERNATIONAL FINANCE CORPORATION (1998), *Emerging Stock Markets Factbook*, World Bank, Washington, D.C.

INSTITUTE OF INTERNATIONAL FINANCE (2005), "Capital Flows to Emerging Market Economies", 24 September 2005, Washington, D.C.

INWENT-MUNICH CENTRE FOR ECONOMIC, ENVIROMENTAL AND SOCIAL STATISTICS (2002), Report of the Workshop "Measuring Democracy & Good Governance", 21.-23 January 2002, Munich, available online at: www.inwent.org/imperia/md/content/bereich4-intranet/abteilung4-06/d1.pdf

INWENT-MUNICH CENTRE FOR ECONOMIC, ENVIROMENTAL AND SOCIAL STATISTICS (2004), Invitation to the Course "On Measuring Democracy, Governance and Human Rights", Munich, Germany, available online at: www.inwent.org/imperia/md/content/bereich4-intranet/abteilung4-06/1.pdf

JÜTTNER, D.J. and J. MCCARTHY (2000), "Modelling a Rating Crisis", mimeo, Macquarie University, Sydney.

KAUFMANN, D. (2003), "Rethinking Governance, Empirical Lessons Challenge Orthodoxy", *Macroeconomics 0308007, Economics Working Paper Archive EconWPA*, World Bank, Washington, D.C., available online at: www.worldbank.org/wbi/governance/pdf/rethink_gov_stanford.pdf#search='%E2%80%9CRethinking%20Governance%2C%20Empirical%20Lessons%20Challenge%20Orthodox.

KAUFMANN, D. (2004), "Corruption, Governance and Security: Challenges for the Rich Countries and the World" *in Global Competitiveness Report 2004/2005,* Palgrave Macmillan Ltd., New York, NY.

KAUFMANN, D. and A. KRAAY (2002*a*), "Growth Without Governance", *Economía,* Vol. 3(1).

KAUFMANN, D. and A. KRAAY (2002*b*), "Governance Indicators, Aid Allocation and the Millennium Challenge Account", World Bank, Washington, D.C., available online at: www.worldbank.org/wbi/governance/pdf/mca_summary.pdf

KAUFMANN, D. and A. KRAAY (2003*a*), "Governance and Growth: Causality Which Way?", World Bank, Washington, D.C., available online at: www.worldbank.org/wbi/governance/pdf/growthgov_synth.pdf

ISBN 92-64-02685-1 © OECD 2006

KAUFMANN, D. and A. KRAAY (2003*b*), "Governance Matters III: New Indicators for 1996-2002 and Addressing Methodological Challenges", *World Bank Policy Research Working Paper,* No. 3106, World Bank, Washington, D.C., available online at: www.worldbank.org/wbi/governance/pdf/govmatters3_synth.pdf

KAUFMANN, D. and P.C. VICENTE (2005), "Legal Corruption", mimeo, World Bank, Washington, D.C.

KAUFMANN, D., A. KRAAY and M. MASTRUZZI (2003), "Governance Matters III: Governance Indicators for 1996-2002", *World Bank Policy Research Working Paper*, World Bank, Washington, D.C., available online at: www.worldbank.org/wbi/governance/pubs/govmatters3.html

KAUFMANN, D., A. KRAAY and M. MASTRUZZI (2005*a*), "Governance Matters IV: Governance Indicators for 1996-2004", *World Bank Policy Research Working Paper*, No. 3630. World Bank Institute., Washington, D.C.

KAUFMANN, D., A. KRAAY and M. MASTRUZZI (2005*b*), "Illustrative Composite Governance Indicator Worldwide Map", World Bank, Washington, D.C., available online at: http://info.worldbank.org/governance/kkz2004/maps.html

KAUFMANN, D., A. KRAAY and P. ZOIDO-LOBATÓN (1999*a*), "Aggregating Governance Indicators", *World Bank Policy Research Working Paper*, No. 2195, World Bank, Washington, D.C., available online at: www.worldbank.org/wbi/governance/pubs/aggindicators.html

KAUFMANN, D., A. KRAAY and P. ZOIDO-LOBATÓN (1999*b*), "Governance Matters", *World Bank Policy Research Working Paper*, No. 2196. World Bank Institute., Washington, D.C., available online at: www.worldbank.org/wbi/governance/pubs/govmatters.html

KAUFMANN, D., A. KRAAY and P. ZOIDO-LOBATÓN (2002), "Governance Matters II: Updated Indicators for 2000/01", *World Bank Policy Research Working Paper*, No. 2771. World Bank, Washington, D.C.

KNACK, S. (2002), "Governance and Growth: Measurement and Evidence" Forum Series on the Role of Institutions in Promoting Economic Growth, USAID, Washington, D.C.

KNACK, S. and M. KUGLER (2002), "Constructing an Index of Objective Indicators of Good Governance", *PREM Public Sector Group*, World Bank, Washington, D.C., available online at: www1.worldbank.org/publicsector/anticorrupt/FlagshipCourse2003/SecondGenerationIndicators.pdf

KNACK, S., M. KUGLER and N. MANNING (2003), "Second-Generation Governance Indicators", *International Review of Administrative Sciences*, Vol. 69.

KRAAY, A. and V. NEHRU (2004), "When is External Debt Sustainable?", *World Bank Policy Research Working Paper*, No. 3200, World Bank Institute., Washington, D.C., available online at: http://www.worldbank.org/hipc/Kraay_Nehru2004.pdf

 ISBN 92-64-02685-1 © OECD 2006

LANDMAN, T. and J. HÄUSERMANN (2003), "Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance", Report for the Statistical Office of the Commission of the European Communities (EUROSTAT), Human Rights Centre, University of Essex.

LEITE, C. and J. WEIDMANN (1999), "Does Mother Nature Corrupt? Natural Resources, Corruption and Economic Growth", *International Monetary Fund Working Paper*, WP/99/85, International Monetary Fund, Washington, D.C.

LIPSET, S. (1959), "Some Social Requisites of Democracy, Economic Development and Political Legitimacy", *American Political Science Review,* No. 53.

LORA, E. and U. PANIZZA (2002), "Structural Reforms in Latin America Under Scrutiny", Text for Seminar "Reforming Reforms", 11 March 2002, Fortaleza, Brazil.

MADDISON, A. (2001), *The World Economy: A Millennial Perspective*, Development Centre Studies, OECD, Paris.

MADDISON, A. (2003), *The World Economy: Historical Statistics*, Development Centre Studies, OECD, Paris.

MARSHALL, M.G. and K. JAGGERS (2002) "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002. Dataset Users Manual", Center for International Development and Conflict Management, University of Maryland, available online at: http://www.cidcm.umd.edu/inscr/polity/

MARSHALL, M.G., K. JAGGERS and T.R. GURR (2005), "Polity IV Project", Data available online at: http://www.cidcm.umd.edu/inscr/polity/index.htm

MEISEL, N. (2004), *Governance Culture and Development: A Different Perspective on Corporate Governance*, Development Centre Studies, OECD, Paris.

METAGORA (2006), "Measuring Human Rights and Democratic Governance to inform key policies: light version of the synthesis report on results and lessons from the two-year pilot phase of the Metagora project", advanced draft.

MEUNIER, N. and T. SOLLOQOUB (2005), *Economie du risque pays*, La Découverte, Paris.

MUNCK, G.L. and J. VERKUILEN (2002), "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices", *Comparative Political Studies*, Vol. 35(1).

NARDO, M., M. SAISANA, A. SALTELLI, S. TARANTOLA, A. HOFFMAN and E. GIOVANNI (2005), *Handbook on Constructing Composite Indicators: Methodology and User Guide*, OECD, Paris, available online at: http://www.olis.oecd.org/olis/2005doc.nsf/LinkTo/std-doc(2005)3

NORTH, D.C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge, United Kingdom.

NORTH, D.C. (2005), *Understanding the Process of Economic Change,* Princeton University Press, New Jersey.

ISBN 92-64-02685-1 © OECD 2006

OECD Development Co-operation Directorate, "Senior Level Forum on Development Effectiveness in Fragile States, Aid Allocation Criteria: Managing for Development Results and Difficult Partnerships", *Oxford Policy Management*, 13-14 January 2005.

OECD (2005), *Management in Government: Comparative Country Data*, OECD Governance Directorate, available online at: http://www.oecd.org/gov/indicators

Olken, B. (2006), "Corruption Perceptions vs. Corruption Reality", *Working Paper*, Harvard University and NBER.

Olson, M. (1965), *The Logic of Collective Action: Public Goods and the Theory of Groups*, Harvard University Press, Cambridge.

Olson, M. (1982), *The Rise and Decline of Nations: Economic Growth, Stagflation and Social Rigidities*, Yale University Press, New Haven.

Olson, M. (2000), *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*, Basic Books, New York, NY.

Oman, C. (2000), *Policy Competition for Foreign Direct Investment: A Study of Competition Among Governments to Attract FDI*, Development Centre Studies, OECD, Paris

Oman, C. and G. Wignaraja (1991), *The Postwar Evolution of Development Thinking*, Palgrave MacMillan Ltd. and Development Centre Study, London.

Ortiz, G. (2003), "Latin America and the Washington Consensus: Overcoming Reform Fatigue", *Finance & Development*, Vol. Sept. 2003.

Przeworski, A., M.E. Alvarez, J.A. Cheibub and F. Limongi (2000), *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*, Cambridge University Press, Cambridge, United Kingdom.

Reisen, H. (2003), "Ratings Since the Asian Crisis", *Development Centre Working Paper* No. 214, OECD, Paris, available online at: www.oecd.org/dataoecd/22/8/1934625.pdf

Reisen, H. (2000), "Basel II: a Risky Strategy", *OECD Observer*, OECD, Paris.

Rigobon, R. and D. Rodrik (2004), "Rule of Law, Democracy, Openness, and Income: estimating the interrelationships", *NBER Working Paper*, No. 10750.

Roodman, D. (2005), "An Index of Donor Performance", *Working Paper*, Center for Global Development, No. 67, Washington, D.C.

Shirley, M. (2002), *on* Knack, S., "Governance and Growth: Measurement and Evidence", USAID, Washington, D.C.

Standard and Poor's (2003), *Global Stock Markets Factbook*, New York, NY.

Sudders, M. and J. Nahem (2004), *Governance Indicators: A User's Guide*, UNDP and EUROSTAT.

The Economist (2005), "A Regime Changes", 2 June.

 ISBN 92-64-02685-1 © OECD 2006

THOMPSON, T. and A. SHAH (2004), "Transparency International's Corruption Perceptions Index: Whose Perceptions Are They Anyway?", Operations Evaluations Department, World Bank, Washington, D.C., unpublished.

TRANSPARENCY INTERNATIONAL (2003), *Frequently Asked Questions About the TI Corruption Perceptions Index (CPI) 2003,* available online at: http://ww1.transparency.org/cpi/2003/cpi2003_faq.en.html

TRANSPARENCY INTERNATIONAL (2004*a*), *Frequently Asked Questions About the TI Corruption Perceptions Index (CPI) 2004,* available online at: www.transparency.org/cpi/2004/cpi2004_faq.en.html#five

TRANSPARENCY INTERNATIONAL (2004*b*), *Press Release Corruption Perceptions Index 2004,* available online at: http://ww1.transparency.org/cpi/2004/cpi2004_faq.en.html

UBS (2005), *Handbook 2004/2005*, Zürich, available online at: www.ubs.com/1/e/investors/annual_reporting2004/handbook.html

UNDP (2002), *Arab Human Development Report 2002*, Geneva.

UNDP (2003), "Overview: Existing Framework of Governance Indicators", available online at: www.undp.org/oslocentre/docsjuly03/Overview%20Existing%20Framework%20of%20Indicators.xls

UNDP (2004), "Sources for Democratic Governance Indicators", Geneva, available online at: www.undp.org/oslocentre/docs04/Indicator%20Sources.pdf

UNDP (2005*a*), *Arab Human Development Report 2004: Towards Freedom in the Arab World*, Oslo Governance Centre, Oslo, available online at: www.undp.org/oslocentre/docs05/Update_april%202005.pdf

UNDP (2005*b*), "Empowering and Engendering Governance Indicators (2005)", UNDP-ICSSR Technical Workshop on Governance Indicators for Pro-Poor and Gender-Sensitive Policy Reform, New Delhi, 20-22 April 2005.

USAID (2005), *Report on the Criteria and Methodology for Determining the Eligibility of Candidate Countries for Millennium Challenge Account Assistance in FY 2005*, Washington, D.C., available online at: www.mca.gov/about_us/congressional_reports /Report%20to%20Congress%20on%20Criteria%20and%20Methodology%20FY051.pdf

VAN DE WALLE, S. (2005), "Measuring Bureaucratic Quality in Governance indicators", Paper prepared for the 8th Public Management Conference, Los Angeles, Sept 29 -Oct 1, Public Management Institute, Leuven, Netherlands.

VAN DE WALLE, S. (forthcoming, 2006), "The State of the World's Bureaucracies", *Journal of Comparative Policy Analysis*.

WILLIAMSON, H. (2004), "Hazards of Charting Corruption", *Financial Times*, 20 October 2004, London.

ISBN 92-64-02685-1 © OECD 2006

WILLIAMSON, J. (2000), "What Should the World Bank Think About the Washington, Consensus?", *World Bank Research Observer*, Vol. 15.

WOOLDRIDGE, J.M. (2001), "Applications of Generalized Method of Moments Estimation", *Journal of Economic Perspectives*, Vol. 15(4).

WORLD BANK (2001), "OED IDA Review: Review of the Performance-Based Allocation System, IDA 10-12", February 14, 2001, available online at: wbln0018.worldbank.org/oed/oeddoclib.nsf/View+to+Link+WebPages/310529285EC4A83685256A680079BC9F/$FILE/PBA.pdf

WORLD BANK (2004*a*), *World Development Report 2005: A Better Investment Climate for Everyone*, Washington, D.C.

WORLD BANK (2004*b*), *World Bank Annual Report 2004*, Washington, D.C.

WORLD BANK (2004*c*), "Operations Policy and Country Services (2004)", Country Policy and Institutional Assessments: 2004 Assessment Questionnaire, Washington, D.C., available online at: siteresources.worldbank.org/IDA/Resources/CPIA2004questionnaire.pdf

WORLD BANK (2005), *Global Development Finance 2005*, World Bank, Washington, D.C.

WORLD BANK (2006), *Strengthening Mutual Accountability – Aid, Trade, and Governance*, World Bank, Washington, D.C.

WORLD BANK INSTITUTE, *Consolidated Matrix of External and Internal World Bank Governance Data Sets and Instruments*, Washington, D.C., available online at: www.worldbank.org/wbi/governance/govdatasets/

WORLD BANK INSTITUTE (2006), *Governance Data: Web-Interactive Inventory of Datasets and Empirical Tools*, Washington, D.C., available online at: www.worldbank.org/wbi/governance/govdatasets/index.html

ISBN 92-64-02685-1 © OECD 2006

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(41 2006 04 1 P) ISBN 92-64-02685-1 – No. 55225 2006

Development Centre Studies

Uses and Abuses of Governance Indicators

Rapidly rising attention to the quality of governance in developing countries is driving explosive growth in the use of governance "indicators" by international investors, donors of official development assistance, development analysts and academics.

This study helps both users and producers of governance indicators to understand the strengths and weaknesses of the best and most widely used indicators, helps them find their way through the jungle of hundreds of existing governance indicator datasets, and shows how governance indicators tend to be widely misused both in international comparisons and in tracking changes in the quality of governance in individual countries. It also explains recent developments in the supply of governance indicators, arguing 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.

Highly informative and equally persuasive.

Adam Przeworski,
Carroll and Milton Petrie Professor of Politics, New York University

Should be required reading by all who publish or use governance indicators, especially those who are making policy or offering policy advice.

John D. Sullivan,
Executive Director, Center for International Private Enterprise

Given the proliferation of governance indicators, Arndt and Oman offer a welcome assessment of how existing studies are best interpreted and used by scholars, aid agencies, governments and businesses. This careful appraisal of present knowledge will be the basis for launching the next round of inquiries.

Hilton L. Root,
Former US Treasury Department official, author of *Capital and Collusion*, Princeton University Press.

A seminal study. Extremely thorough. Should reach a very wide audience.

François Roubaud,
Director of Research, *Développement, Institutions et Analyses de Long terme* (DIAL)

The full text of this book is available on line via these links:
http://www.sourceoecd.org/development/9264026851
http://www.sourceoecd.org/emergingeconomies/9264026851
http://www.sourceoecd.org/governance/9264026851

Those with access to all OECD books on line should use this link:
http://www.sourceoecd.org/9264026851

SourceOECD is the OECD's online library of books, periodicals and statistical databases. For more information about this award-winning service and free trials ask your librarian, or write to us at **SourceOECD@oecd.org**.

This work is published under the auspices of the OECD Development Centre. The Centre promotes comparative development analysis and policy dialogue, as described at:

www.oecd.org/dev

ISBN 92-64-02685-1
41 2006 04 1 P

www.oecd.org

OECD PUBLISHING