

Governance indicators

Citation for published version (APA):

Arndt, C. E. (2009). *Governance indicators*. [Doctoral Thesis, Maastricht University]. Boekenplan.
<https://doi.org/10.26481/dis.20090605ca>

Document status and date:

Published: 01/01/2009

DOI:

[10.26481/dis.20090605ca](https://doi.org/10.26481/dis.20090605ca)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Governance Indicators

©2009 Christiane Arndt

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form, or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing, from the author.

ISBN: 978-9-08-666103-9

Publisher: Boekenplan, Maastricht

GOVERNANCE INDICATORS

DISSERTATION

to obtain the degree of Doctor
at the Maastricht University,
on the authority of the Rector Magnificus
Prof. Dr. G.P.M.F. Mols
in accordance with the decision of the Board of Deans
to be defended in public
on Friday, 5 June 2009

by
Christiane Elisabeth Arndt

Supervisor:

Prof. Dr. C.J.N. de Neubourg

Assessment Committee:

Prof. Dr. C. Hood
Prof. Dr. A. Szirmai
Dr. S. Van de Walle
Prof. Dr. P. Van den Bossche
Prof. Dr. B. Verspagen

Acknowledgements

Completing a PhD is a marathon event, and I would not have been able to complete this journey without the aid and support of countless people over the past years. I first would like to express my gratitude towards my advisor, Professor Chris de Neubourg. His advice, encouragement and humour have been invaluable to my research.

The Maastricht Graduate School of Governance is a wonderful place. I could not have imagined a better place for completing my dissertation. I would like to thank my fellow PhD students. They each helped me make my time in the PhD program more fun and interesting. I would also like to thank the staff of the Governance School for their support and for the smiles on their faces and their encouragement. Furthermore, I am grateful to the participants of the May 2006 Maastricht Graduate School of Governance seminar for their encouragement and comments. I would also like to thank the Graduate School of Governance for their financial support.

My dissertation would not be as it is today without the continuous advice from Denis de Crombrughe. His econometric and statistical knowledge as well as his analytical and pedagogical skills were invaluable to my research. From the beginning, he was always there for me. I will miss the long and enriching discussions with him.

I would like to thank the OECD Development Centre for the enriching years of research I spent there. In particular, I would like to thank Charles Oman with whom I worked and published together all those years and from whom I learned a lot, both in terms of substance and writing. I am also very grateful to Orsetta Causa and Ekkehard Ernst, Thomas Heimgartner, Nico Jaspers, Louka Katseli, Juan Raomón de Laiglesi, Jana Malinska, Nick Manning, Nicolas Meisel and Helmut Reisen, who were my colleagues at that time at the OECD and provided me with excellent feedback and advice. I thank the participants in the November 2005 informal Development Centre Seminar for their very useful comments. I would also like to thank Meria Puhakka who did an exceptional job in promoting the book “Uses and Abuses of Governance Indicators”. The resulting publicity led to many enriching contacts which were very beneficial to my

research. I would also like to acknowledge 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 for the project on Governance Indicators and on Governance and Investment.

I am very grateful to the Malcolm Wiener Center for Social Policy at the John F. Kennedy School of Government at Harvard University where I spent the year 2007 as a research fellow. Kati Kriz welcomed me very warmly in Harvard and ensured that the year was a truly enriching experience for me. My research benefited from enormously valuable suggestions and comments from Matthew Andrews, Ricardo Hausmann, Pippa Norris, Lant Pritchett, Dani Rodrik and James Stock at Harvard University. I am very grateful to the participants of the seminar on Governance Indicators at the Kennedy School of Government for their extremely useful input. I would also like to thank the Harvard Shuttle Team, in particular Jim, Marc, Bob and Bonny, who ensured that I could reach the library and my office despite a foot injury I had during that time.

I am very grateful to the experts at the World Bank that I met over the years and who shared their rich experiences and knowledge with me. In particular, I would like to thank Stephen Knack and Nick Manning for the enriching discussions and their feedback on my work. I would also like to thank Daniel Kaufmann and Aart Kraay at the World Bank Institute for their frank and friendly critical feedback. I thank James Anderson and Jeffrey Rinne for their extremely helpful comments. I would also like to thank the experts who I interviewed for my research and who so generously shared with me their insights.

Over the course of the years, many people commented on papers and book chapters that were part of my dissertation. My research benefited from enormously useful suggestions and comments from Christopher Hood and Ruth Dixon at Oxford University, Jacques Ould-Aoudia from France's Ministry of Finance and Economy, Javier Herrera, Mireille Razafindrakato and Francois Roubaud from DIAL, Guilhem Bascle at HEC Paris and Simon Kaja at the University of British Columbia. I would further like to express my gratitude to the participants of the workshop "Ranking and rating public services" in Oxford, and to those of the seminar "The Politics of Governance

Ratings” in Syracuse for their extremely thorough comments and suggestions for improvement. My research benefited from the very constructive discussions and the provision of useful material by Marie Wolkers and the team at Transparency International. I would also like to thank Michaela Krause from Tilburg University, Richard Öhrvall from Statistics Sweden, Jan Robert Suesser from France's Ministry of Finance and Economy and John Sullivan from CIPE for their rich comments. I am grateful to Nicolas Meunier for sharing his technical expertise on financial markets and country risk analysis. Furthermore, I am very grateful to the members of the Reading Committee of the OECD Development Centre for their excellent comments and suggestions. My research also benefited from the very helpful input of two anonymous reviewers of the Journal of Public management. Finally, I would like to warmly thank the members of my dissertation committee, Christopher Hood, Adam Szirmai, Steven Van de Walle, Peter Van den Bossche and Bart Verspagen for their insightful comments and advice.

My research bases many of its findings on interviews with companies, banks, donors of official development assistance, risk analysts, academics and OECD and World Bank staff who requested anonymity but whose assistance was invaluable.

I would like to thank Moritz Neuffer for his thorough rereading of my thesis, his thoughtful and precise comments and that I could always count on him. I thank Manuel Müllers for his continuous support and patience during the last phase of the thesis and for his excellent assistance, together with Jessica Schmitz, with the layout. I thank Sabine Filla, Ana Maria Daza Vargas and Tobias Baur for their emergency help.

I am very grateful to my friends for their support and patience. In particular, I would like to thank Sabine Filla, who was always there for me. I thank my flatmates Ana Maria Daza Vargas, Florian Henning, Metka Hercog and Melissa Siegel for our wonderful dinners and wine sessions. I would also like to thank Victor Cebotari, Jinjing Li, Denisa Sologan and Hao Yu for all those late-night sessions in the PhD kitchen. I would like to thank Françoise Loiret for the enriching discussions and her advice. I am also grateful to my friends Samik Basu, Laura Georg, Isa Gras, Rudolf Hömberg, Simon Kaja, Lutz

Krebs and Pia Streicher who shared the experience of writing a dissertation with me. I am also thankful to the authors of the PhD comic book.

Finally, I am very grateful to my parents, my sister Regina and my brother Martin for all their support, understanding and love. I am thankful to Guilhem for his advice and for those special and joyful moments that gave me the energy for my research.

Contents

List of boxes	XIV
List of figures	XVI
List of tables	XVIII
1 Introduction	21
1.1 Why all the interest in governance?	24
1. 1. 1 International investment	24
1. 1. 2 End of the Cold War	26
1. 1. 3 Failed policy reform	27
1. 1. 4 New Institutional Economics	29
1.2 Sources of governance indicators	29
1.3 Uses of governance indicators	30
1.4 The Worldwide Governance Indicators	31
1.5 The politics of governance ratings	32
1.6 The Rule of Law and economic growth: A panel data re-examination	33
1.7 Outline	34
2 Sources of governance indicators	37
2.1 International Country Risk Guide	38
2.2 Freedom House	41
2.3 Transparency International	45
2.4 The World Bank	48
2.4.1 International Development Association	48
2.4.2 World Bank Institute	50
2.5 Finding your way through the jungle of governance indicators	54
3 Uses of governance indicators	59
3.1 International Investors	60
3.2 Donors of aid	67
3.2.1 World Bank CPIAs	69
3.2.2 Transparency International's CPI	70

3.2.3 Millennium Challenge Account's use of KKZ indicators	72
3.3 Analysts and academics	74
3.3.1 CPIAs	75
3.3.2 CPI and FH indicators	76
3.3.3 KKZ indicators	77
4 The Worldwide Governance Indicators	83
4.1 Introduction	84
4.2 History, the underlying aggregation model and characteristics of the WGIs	85
4.2.1. History and overview	85
4.2.2 Model and assumptions	89
4.2.3 Characteristics of the WGIs	110
4.3 Concerns	116
4.3.1 Correlation of disturbances	117
4.3.2 Lack of transparency	125
4.3.3 Weak conceptual foundation	127
4.3.4 Sample bias and weighting method	130
4.3.5 Lack of comparability over time	142
4.4 Conclusion	151
5 The politics of governance ratings	153
5.1 World Bank governance indicators	156
5.1.1 The Worldwide Governance Indicators	158
5.1.2 The Doing Business Indicators	159
5.1.3 Other World Bank indicators	160
5.1.4 Move towards "actionable" indicators	162
5.2 Why does the World Bank produce cross-country comparable governance indicators? Do their indicators satisfy the purposes they are produced for?	164
5.2.1 Help identify areas of reform and measure the success of governance reform	165
5.2.2 Incentives for developing countries to improve their governance and agenda-setting	170
5.2.3 Enhance country dialogue	172
5.2.4 Higher transparency of aid-allocation decisions	174
5.2.5 Quantitative governance analysis	175

5.2.6 Institutional reasons	176
5.3 An attempt to explain the dominance of the Worldwide Governance Indicators over more meaningful indicators	179
5.3.1 Incomplete understanding of indicators' limitations	180
5.3.2 Supply meets demand for summary measurements with broad country coverage	182
5.3.3 Media	185
5.3.4 Herd behaviour	186
5.3.5 Perceived absence of an alternative	187
5.4 Conclusion and recommendations	187
6 Rule of Law and economic growth:	
A panel data re-examination	193
6.1 Growth model	201
6.2 Data	209
6.2.1 Data periodicity and scaling	209
6.2.2 Rule of Law	210
6.2.3 Dependent variable and other explanatory variables	221
6.3 Specification issues and estimation methods	226
6.4 Empirical results	230
6.4.1 Growth equations without governance indicators	230
6.4.2 Introducing the Rule of Law indicators	232
6.4.3 Fragility of significance of other regressors	236
6.4.4 Rule of Law and the investment rate	238
6.5 Conclusion	239
Appendix Chapter 6: detailed tables	250
7 Conclusion	273
7. 1 Sources of governance indicators (chapter two)	274
7.2 Uses of governance indicators (chapter three)	276
7. 3 The Worldwide Governance Indicators (chapter four)	278
7.4 Politics of governance ratings (chapter five)	280
7.5 Rule of Law and economic growth: A panel data re-examination (chapter six)	285
7.6 The governance indicator paradox	287

7.7 Limitations and recommendations for future work	287
Reference list	289
Appendix: Interview methodology	309
I. Introduction	310
I.1 Interviews of corporate and financial investors	311
I.2 Bilateral donors	314
I.3 World Bank staff	315
I.4 Staff of other international organizations	315
I.5 Academics	316
Summary in Dutch/Samenvatting	317
Curriculum Vitae	325

List of boxes

Box 2.1. Governance Indicator Guides and Inventories	57
Box 4.1. Sources of Governance Data	88
Box 4.2 “Regulatory Quality”	103

List of figures

Figure 4.1 Bangladesh Regulatory Quality 1998	111
Bangladesh Regulatory Quality 2005	112
Figure 4.2 Weights 2004, CC	113
Figure 4.3 Weights over time, CC	114
Figure 4.4 Regulatory Quality (2006)	139
Figure 4.5 a) "Regulatory Quality" 1998 and 2005 in El Salvador and Bolivia, "Governance Matters V"	148
Figure 4.5 b) "Regulatory Quality" 1998 and 2005 in El Salvador and Bolivia, "Governance Matters VI"	148

List of tables

Table 4.1 Weights 2006 by Source and Worldwide Governance Indicator	108
Table 4.2 "Regulatory Quality" 2006 in Costa Rica, Guatemala, Honduras, Nicaragua and Panama	140
Table 4.3 a) Sources for "Regulatory Quality" 1998 El Salvador and Bolivia in "Governance Matters V" Data for figure 4.5a)	149
Table 4.3 b) Sources for "Regulatory Quality" 1998 El Salvador and Bolivia in "Governance Matters VI" Data for figure 4.5b)	150
Table 6.1 Rule of Law indicators	214
Table 6.2 Correlations between different Rule of Law indicators	219
Table 6.3 Strengths and Weaknesses of Rule of Law indicators	220
Table 6.4 Correlations between Rule of Law indicators and regime type, income and growth	222
Table 6.5 Dependent variable and main control variables	224
Table 6.6: Summary table of estimation methods	229
Table 6.7(a): Different methods, full samples, ordinary SEs	242
Table 6.7(b): Different methods, same samples, ordinary SEs	243
Table 6.8(a): Instrumenting income, full samples	244
Table 6.8(b): Instrumenting income, same samples	245
Table 6.9: Pseudo-metric Rule of Law indicators	246
Table 6.10: Categorical Rule of Law dummies	247
Table 6.11: Categorical Rule of Law dummies, instrumenting Income	248

Table 6.12: Categorical Rule of Law dummies in the investment equation	249
Table 6.A.01(a): Pooled OLS, full samples	250
Table 6.A.01(b): Pooled OLS, same samples	251
Table 6.A.02(a): RE1 (Time Averages of control variables), full samples	252
Table 6.A.02(b): RE1 (Time Averages of control variables), same samples	253
Table 6.A.03(a): RE2 full samples	254
Table 6.A.03(b): RE2 same samples	255
Table 6.A.04(a): FE full samples	256
Table 6.A.04(b): FE same samples	257
Table 6.A.05(a): Pooled IV, full samples	258
Table 6.A.05(b): Pooled IV, same samples	259
Table 6.A.06(a): RE-IV-TAv (Time Averages of control variables), full samples	260
Table 6.A.06(b): RE-IV-TAv (Time Averages of control variables), same samples	261
Table 6.A.07(a): RE2-IV full samples	262
Table 6.A.07(b): RE2-IV same samples	263
Table 6.A.08(a): FE-IV full samples	264
Table 6.A.08(b): FE-IV same samples	265
Table 6.B.01(a): Investment rate, Pooled OLS, full samples	266
Table 6.B.01(b): Investment rate, Pooled OLS, same samples	267
Table 6.B.02(a): Investment rate, RE full samples with time averages of control variables	268
Table 6.B.02(b): Investment rate, RE same samples with time averages of control variables	269
Table 6.B.03(a): Investment rate, RE2 full samples	270

Table 6.B.03(b): Investment rate, RE2 same samples	270
Table 6.B.04(a): Investment rate, FE full samples	271
Table 6.B.04(b): Investment rate, FE same samples	271

1 Introduction

Over the last decade the quality of governance in developing and emerging-market economies has moved into the spotlight of development economists, international investors, bi- and multilateral development co-operation agencies and the media. Following the maxim that you can only manage what you can measure, they have turned widely to using quantitative indicators of the quality of governance for analysis and decision-making, with far-reaching consequences for developing countries.

Given that the use of governance indicators is a relatively recent phenomenon, a comprehensive and in-depth analysis of the actual use and (mis)use of governance indicators and the strengths and weaknesses of the most popular governance indicators was missing in the literature.

Starting from what is known from previous publications, this study:

- Identifies the most popular governance indicators, explains their content and puts them into the historical context (chapter 2);
- Identifies the major users of governance indicators and analyses their use of indicators (chapter 3);
- Provides an in-depth analysis of the most popular governance indicators, the Worldwide Governance Indicators of the World Bank Institute (chapter 4);

- Analyses the politics of governance indicators, new trends in governance indicators and gives recommendations for users and producers of indicators (chapter 5);
- Investigates, using panel data, the relationship between governance and economic growth and tests its fragility with respect to the specific choice of Rule of Law indicators, as a major subset of governance indicators (chapter 6).

Chapters two and three have been published in a book¹. The last three chapters are self-contained papers. Because of the fact that the chapters in this dissertation have been or will be published as separate papers, a small overlap especially in the introductory sections is inevitable.

The next section briefly explains the rise in interest in the quality of governance and proceeds then to explain the content, methodology and outline of each chapter as well as the contribution it makes to policy-making and academic research.

¹ Arndt and Oman (2006).

1. 1 Why all the interest in governance?²

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.

1. 1. 1 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 \$100 billion since 1996 and over \$200 billion since 2004. Equally spectacular and important for driving up interest in the quality of local governance has been the growth of international *portfolio* investment in developing and “emerging market”³ economies — notably by major pension funds and other large institutional investors — from net annual flows of below \$300 million in portfolio equity to about \$16 billion in the

² The section “Why all the interest in governance” was published as chapter 1 in Arndt and Oman (2006).

³ 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.

1990s and \$50 billion in 2000-2007, and another \$30 billion in net annual flows in bonds in 2000-2007.⁴

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)⁵.

⁴ World Development Indicators Online 2008 (World Bank, 2008a).

⁵ 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."

1. 1. 2 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 D. 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 1980s⁶. OECD governments, their national aid agencies and multilateral development organisations (including the OECD 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

⁶ The remark was made, of course, prior to the onset of the Cold War.

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 management fell by 49 per cent from 2003 to 2008, its lending to improve public sector governance rose by 76 per cent during that period⁷.

1. 1. 3 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

⁷ World Bank Annual Report 2008.

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).

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 reforms⁸ 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

⁸ See, for example, Ortiz (2003), Lora and Panizza (2002).

local governance may go far to explain the relative reform failures of the 1980s and 1990s⁹.

1.1.4 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).

1. 2 Sources of governance indicators

While there are hundreds of governance indicator datasets available, a few dominate the market. Investors, donors, journalists and development analysts rely very heavily on a small selection of

⁹ 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).

composite, mostly perception-based indicators. Chapter two provides an overview of the “Governance Indicators Market”, useful for researchers and practitioners. It is based on literature research and interviews¹⁰.

Chapter two starts with introducing five of the most widely-used governance indicators and puts them into the historical context. It proceeds then to classify indicators into perceptions-based and facts-based indicators and explains why perceptions-based indicators are more popular. It ends with an overview of governance inventories helpful to find the right indicator for every purpose and to understand their strength and weaknesses.

1.3 Uses of governance indicators

The supply of governance indicators is booming and continues to grow. Who uses these indicators for what purpose? Are indicators correctly used? What are the consequences for developing countries? Chapter three fills a gap in the literature by providing a comprehensive analysis of the actual use of governance indicators. It shows that ratings can have far-reaching implications for developing countries, in terms of resource allocation and reputation. It also shows frequent misuse of indicators. This chapter is directly relevant for a wide-range of audiences, including development economists and

¹⁰ See appendix for a description of the methodology.

decision-makers in bi- and multilateral organisations. Findings are based on surveys and personal interviews with the management and staff of ten major corporate and financial investors, as well as with spokespersons for business associations and official development co-operation agencies in Europe and the United States¹¹.

Chapter three starts with an analysis of the use of indicators by investors, proceeds to analyse the use of indicators by donors of aid and finally analyses the use of indicators by development analysts and academics. It points out incorrect use of indicators in each section.

1.4 The Worldwide Governance Indicators

As the preceding chapter showed, governance indicators are both widely used and misused, with far-reaching consequences for developing countries. Given the rising importance of governance indicators, this chapter therefore provides an in-depth analysis of the most popular governance indicators, the Worldwide Governance Indicators (WGIs) of the World Bank Institute. Users of these indicators often fail to take into account the limitations of the WGIs and therefore overstretch the use of the WGIs to purposes they should not be used for.

¹¹ See appendix for a description of the methodology.

Chapter four starts with an analysis of the history, the underlying aggregation model and characteristics of the WGIs and proceeds then to discuss five interrelated problems with the WGIs with direct relevance for users of the indicators as well as for future work on indicators.

1.5 The politics of governance ratings

The World Bank plays a leading role in the supply of cross-country comparable governance ratings. Chapter five draws on interview of World Bank staff and other international organisations¹² to understand why the World Bank produces cross-country comparable indicator to bring to light the current controversies within the World Bank about existing indicators and future work on governance indicators. The outcome of this debate will heavily impact World Bank policies in developing countries, not the least because indicators determine aid-allocation and are expected to play a crucial part in setting the reform agenda in developing countries and assessing the success of reforms undertaken. The chapter also attempts to explain why so many external users rely on the World Governance Indicators, despite the limitations of these indicators, the reluctance within the World Bank to use them and the availability of more meaningful indicators. Finally, it gives recommendations to both consumers and producers of indicators for the short- and long-run.

¹² See appendix for a description of the methodology.

Chapter five first places its contribution into existing literature. It then introduces the governance indicators produced at the World Bank and explains the motivation within the World Bank to produce governance indicators and whether they satisfy the expectations. In a next step, it explains why use of the WGIs is so widespread, in particular outside the World Bank. It concludes with a set of recommendations.

1.6 The Rule of Law and economic growth: A panel data re-examination

The reorientation in development economics and policy towards governance has been backed-up and partly based on quantitative studies that find a positive impact of good governance on economic growth. Most of this literature bases its findings on cross-country data and on aggregate income-level data. Moreover, there are few studies on the impact of good governance on growth over a policy-relevant period of ten to thirty years.

This chapter draws on a comprehensive panel data set and a variety of panel data methods with a time-recursivity system to identify the effect of governance on economic growth over the last decades. Given that governance is an encompassing term and that there are hundreds of indicators available, ranging from democracy indicators to corruption and property rights indicators, we choose to focus on one aspect of governance that has been the focus of attention

of development economics over the last decade and has frequently been used as a proxy for governance in growth studies, the “Rule of Law”. According to *The Economist*, the Rule of Law has become “the motherhood and apple pie of development economics” (*The Economist*, 2008). We control for other clearly distinguishable aspects of governance such as regime type. Chapter 6 tests the robustness of econometric results depending on the choice of the proxy indicators for Rule of Law and the time period chosen. It deals with heterogeneity in the form of unobserved random effects, allowing for correlation with some of the regressors in the spirit of Mundlak (1978).

Chapter six first explains the choice of growth model and proceeds to explain the dataset. Subsequently, it discusses findings of pooled cross-section data methods and fixed-effects methods and of random effects methods. Finally, it concludes with implications of the findings for academic analysis and policy-makers including recommendations for future research.

1.7 Outline

Next to this introductory chapter, this dissertation consists of two book chapters (chapters 2 and 3), three papers (chapters 4 to 6) and a concluding chapter (chapter 7).

1. Introduction

The conclusion discusses the main conclusion of each chapter and its implications for academic research and policy-making. It indicates the limitations of the study and suggests areas for further research.

2 Sources of governance indicators^{13, 14}

¹³ Research for chapters two, three and four was undertaken together with Charles Oman at the OECD Development Centre.

¹⁴ This chapter has been published as chapter two in Arndt and Oman (2006). After publication of the book, the World Bank Institute started to publish its indicators on a yearly basis and renamed them “Worldwide Governance Indicators”. The term commonly used before to refer to the World Bank Institute Indicators was “KKZ” (Kaufmann, Kraay, Zoido), which is used throughout this chapter and in chapter 3. Furthermore, after publication of the book, Freedom House decided to disclose countries’ scores on subcategories, but not on each individual indicator.

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.

2.1 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 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

2. Sources of governance indicators

operations”¹⁵. 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);

¹⁵ See <http://www.icrgonline.com> (our emphasis).

- 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.

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.

2. Sources of governance indicators

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

2.2 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”¹⁶. 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”.

¹⁶ <http://www.freedomhouse.org>.

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 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,

2. Sources of governance indicators

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.
- 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 undue 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

2. Sources of governance indicators

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).

2.3 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.

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 communities¹⁷. The CPI attracted massive global attention and helped to put corruption on the global development agenda¹⁸.

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

¹⁷ See <http://www.transparency.org>.

¹⁸ 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" organized 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.

2. Sources of governance indicators

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 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.

2.4 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.

2.4.1 International Development Association

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

2. Sources of governance indicators

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
- 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 results¹⁹.

2.4.2 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

¹⁹ See Gelb, Ngo and Ye (2004) for estimates of the measurement error inherent in the CPIA.

2. Sources of governance indicators

CPI, they have played a leading role in putting governance on the agenda in developing countries²⁰.

Produced by the WBI's Daniel Kaufmann, originally with co-authors Aart Kraay and Pablo Zoido-Lobaton (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 concerns²¹. One was the apparent lack of robustness of cross-country comparisons using different individual data sources, especially when the different sources led to different conclusions. Second was concern about how to interpret cross-country differences and their statistical and practical significance. Third was concern that it is difficult to compare results from regional surveys with broader cross-country surveys. A fourth concern was to find a way to produce useful overarching, integrative or summary indicators, given the large and growing diversity of individual sources cropping up in different pieces of research and in policy debates.

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

²⁰ See press coverage both within and outside developing countries at: <http://www.worldbank.org/wbi/governance/press-2004indicators.html>.

²¹ We thank Aart Kraay for this information

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 made²²).

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

²² See Chapter 4.

2. Sources of governance indicators

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 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 available²³.

²³ 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).

2.5 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.

2. Sources of governance indicators

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 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 information²⁴.

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

²⁴ See Knack, Kugler and Manning (2003) for a discussion of policy-relevant facts-based indicators.

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.

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.

3 Uses of governance indicators²⁵

²⁵ This chapter was published as chapter three in Arndt and Oman (2006). Findings are based on surveys and personal interviews with the management and staff of ten major corporate and financial investors, as well as with spokespersons for business associations and official development co-operation agencies in Europe and the United States.

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.

3.1 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 billion²⁶.

²⁶ 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, September, 24th 2005.

3. *Uses of governance indicators*

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 Dépôts 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 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

3. Uses of governance indicators

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 reserve²⁷, 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

²⁷ 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.

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 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

3. Uses of governance indicators

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' behaviour²⁸. 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 2002²⁹

Dissatisfaction with the traditional ratings systems has greatly reinforced international investors' attention to the quality of governance, and their demand for governance indicators, in

²⁸ 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.

²⁹ Griffith-Jones and Gottschalk (2006).

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 ceilings³⁰. The less developed the economy of a country is, furthermore, the more importance creditors and investors tend to give to these indicators³¹. 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

³⁰ 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).

³¹ 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.

3. *Uses of governance indicators*

developing as opposed to OECD countries. These interviews also highlighted investors' strong tendency to use *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.

3.2 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 criteria³².

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 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

³² The reliability of the Burnside and Dollar study has been heavily criticized, in particular with respect to the econometric support for the claims in the report (see Hermes and Lensink, 2001 for a summary of the criticism).

3. *Uses of governance indicators*

more overall aid and 50 per cent more finance from the World Bank IDA facility...”³³.

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.

3.2.1 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

³³ Burnside and Dollar (2004, p.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)”.

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.

3.2.2 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 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 table³⁴. Yet the CPI is also widely misused.

³⁴ 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

3. *Uses of governance indicators*

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 and Sampford, 2005), it is often (mis)interpreted by newspapers, and sometimes (mis)used by donors, as precisely such a measure³⁵. 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

Korea has even pledged to reach position 10 or above by 2007 — a tall order, as Seoul was ranked 50th last year”.

³⁵ At least one major donor is known to have decided to stop funding a country on the basis of its CPI score, for example.

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 and Sampford, 2005, p.12).

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).

3.2.3 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). 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*,

3. *Uses of governance indicators*

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; Kaufmann and Kraay, 2002a, 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.

3.3 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,

3. Uses of governance indicators

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 and Sampford, 2005, p.15).

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.

3.3.1 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, p.7).

3.3.2 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, p.4).

The missing theory is also reflected in very different interpretations of the same indicator. Different analysts use the same

3. *Uses of governance indicators*

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).

3.3.3 **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.*, 2005b). 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 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 losing 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

3. Uses of governance indicators

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.

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

3. *Uses of governance indicators*

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 6.³⁶

³⁶ Chapter 6 in Arndt and Oman (2006). In this PhD thesis, the trend towards “actionable” governance indicators is discussed in chapter 5 (The Politics of Governance Ratings).

4 The Worldwide Governance Indicators

4.1 Introduction³⁷

The six Worldwide Governance Indicators (WGIs) of the World Bank Institute, often also abbreviated by their original authors Kaufmann, Kraay and Zoido (KKZ), are the most popular governance indicators. Publicly available and easily accessible on the web, they are widely-used to compare the quality of governance over time and across countries, for aid-allocation decisions, for risk ratings, for academic analysis and for media articles. They summarize data from more than 30 expert assessments and household and firm surveys, and are available from 1996 to today for more than 200 countries.

We take an in-depth look at the strength and weaknesses of the WGIs. The first part explains why and how the WGIs were constructed and the second part discusses five concerns with the WGIs, of which users seem to be widely unaware. The five interrelated concerns tackle i) the key assumption of non-correlated disturbances, ii) the transparency of the ratings, iii) the conceptual foundation, iv) sample selection and the weighting method and v) the comparability over time and across countries.

³⁷ This chapter is based on Arndt and Oman (2006), Arndt, Knack and Oman (2007), Arndt and Romero (2008) and invaluable statistical advice from Denis de Crombrughe.

4.2. History, the underlying aggregation model and characteristics of the WGI

The following section looks at the reasons for the World Bank Institute's decision to produce the WGI, the aggregation model that underlies the final country ratings and the characteristics of the WGI.

4.2.1. History and overview

According to one of the authors³⁸, the WGI were created in response to four inter-related concerns: One was the apparent lack of robustness of cross-country comparisons using different individual data sources, especially when the different sources led to different conclusions. Second was concern about how to interpret cross-country differences and their statistical and practical significance. Third was concern that it is difficult to compare results from regional surveys with broader cross-country surveys. A fourth concern was to find a way to produce useful overarching, integrative or summary indicators, given the large and growing diversity of individual sources cropping up in different pieces of research and in policy debates.

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

³⁸ We thank Aart Kraay for this information.

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.

4. The Worldwide Governance Indicators

- 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.
 - *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 more than 30 different data sources (see Box 4.1), supplied by 30 different organizations. 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.

Box 4.1. Sources of Governance Data

used to construct the KKZ indicators

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

4. The Worldwide Governance Indicators

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.
- *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

4. The Worldwide Governance Indicators

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.
- *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

4. The Worldwide Governance Indicators

them avoid or minimise risk and travel disruptions abroad.
Source for: PS.

- *Gray Area DynamicsTM*: 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 DynamicsTM 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.
- *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 in 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.

4. The Worldwide Governance Indicators

- *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.
- *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.

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

Information based on Kaufmann *et al.* (2005b) and sources' websites.

4.2.2 Model and assumptions

The aggregation procedure is based on the following model:

Kaufmann, Kraay and Mastruzzi use an unobserved components model:

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

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.
$\varepsilon(j,k)$	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.
$\alpha(k), \beta(k)$	Coefficients serving to map, together with the disturbance term $\varepsilon(j,k)$, unobserved governance into the observed data.
$\sigma_{\varepsilon}^2(k)$	Variance of the disturbance terms of indicator k common to all countries j

The unobserved components model is used to account for the fact that governance itself is not observable and that it can only be approximated 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 following assumptions underlie 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.

The “true” level of the quality of governance in a specific country is unknown. Kaufmann *et al.* estimate it conditionally on the results they obtain for each indicator in each of the six governance clusters. These indicators are aggregated in order to estimate the “true” level for each governance cluster.

4. The Worldwide Governance Indicators

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:

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

K = number of indicators

J = number of countries

$y(j)$ = the $K \times 1$ vector of the $y(j,k)$'s for country j

y = the $JK \times 1$ vector of the $y(j,k)$'s for all countries

α = $K \times 1$ vector of the $\alpha(k)$'s

β = $K \times 1$ vector of the $\beta(k)$'s

$$\Omega = \beta\beta' + \text{diag} \{ \sigma_\varepsilon^2(k) \bullet \beta(k)^2 \}$$

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:

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

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 $\alpha(k)$ is subtracted from each observed score $y(j,k)$ and the result is then divided by $\beta(k)$. 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, $\varepsilon(j,k)$, 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).

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

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

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 paragraph describes the estimation of the “true” level of governance (Equation 4) and its standard deviation (Equation 5) step by step.

The several hundred indicators from the more than thirty 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. The WGI “Regulatory Quality” will serve as an illustration. Box 4.2 shows in bold all the sources –

4. The Worldwide Governance Indicators

surveys or expert assessments - underlying the WGI “Regulatory Quality. Most sources are themselves composite indicators, constructed from experts’, business’ and households’ answers to survey questions or checklists. To illustrate, KKZ selected eight survey questions from the World Economic Forum’s Executive Opinion survey for their “Regulatory Quality” indicator and four indicators from the Economist Intelligence Unit.

Box 4.2 "Regulatory Quality"

Representative Sources

GLOBAL INSIGHT GLOBAL RISK SERVICE

Regulations -- Exports: A 2% reduction in export volume as a result of a worsening in export regulations or restrictions (such as export limits) during any 12-month period, with respect to the level at the time of the assessment. *Regulations -- Imports:* A 2% reduction in import volume as a result of a worsening in import regulations or restrictions (such as import quotas) during any 12-month period, with respect to the level at the time of the assessment.

Regulations -- Other Business: An increase in other regulatory burdens, with respect to the level at the time of the assessment that reduces total aggregate investment in real LCU terms by 10%

Ownership of Business by Non-Residents: A 1-point increase on a scale from "0" to "10" in legal restrictions on ownership of business by non-residents during any 12-month period.

Ownership of Equities by Non-Residents: A 1-point increase on a scale from "0" to "10" in legal restrictions on ownership of equities by non-residents during any 12-month period.

ECONOMIST INTELLIGENCE UNIT

Unfair competitive practices

Price controls

Discriminatory tariffs

Excessive protections

WORLD ECONOMIC FORUM GLOBAL COMPETITIVENESS REPORT

Administrative regulations are burdensome

Tax system is distortionary

Import barriers as obstacle to growth

Competition in local market is limited

Anti monopoly policy is lax and ineffective

Environmental regulations hurt competitiveness

Complexity of tax System

Easy to start company

HERITAGE FOUNDATION INDEX OF ECONOMIC FREEDOM

Foreign investment

Banking / finance

Wage/Prices

MERCHANT INTERNATIONAL GROUP GRAY AREA DYNAMICS

Unfair Competition.

Unfair Trade.

POLITICAL RISK SERVICES INTERNATIONAL COUNTRY RISK GUIDE

Investment Profile.

GLOBAL INSIGHT BUSINESS CONDITIONS AND RISK INDICATORS

Tax Effectiveness: How efficient the country's tax collection system is.
Legislation: An assessment of whether the necessary business laws are in place.

Non-representative Sources

AFRICAN DEVELOPMENT BANK COUNTRY POLICY AND INSTITUTIONAL ASSESSMENTS

Trade policy
Competitive environment
Labour Market Policies

ASIAN DEVELOPMENT BANK COUNTRY POLICY AND INSTITUTIONAL ASSESSMENTS

Trade Policy and Forex Regime
Enabling Environment for Private Sector Development
BUSINESS ENTERPRISE ENVIRONMENT SURVEY

Information on the laws and regulations is easy to obtain
How problematic are anti competitive practices for the growth of your business.
How problematic are unpredictable regulations for the growth of your business.
How problematic are labour regulations for the growth of your business.
How problematic are tax regulations for the growth of your business.
How problematic are custom and trade regulations for the growth of your business.

BERTELSMANN TRANSFORMATION INDEX

Competition
Price Stability

WORLD BANK COUNTRY POLICY AND INSTITUTIONAL ASSESSMENTS

Competitive environment
Trade policy

EUROPEAN BANK FOR RECONSTRUCTION AND DEVELOPMENT TRANSITION REPORT

Price liberalization
Trade & foreign exchange system
Competition policy

IFAD RURAL SECTOR PERFORMANCE ASSESSMENTS

Enabling conditions for rural financial services development
Investment climate for rural businesses

Access to agricultural input and produce markets

INSTITUTE FOR MANAGEMENT AND DEVELOPMENT WORLD COMPETITIVENESS YEARBOOK

Access to capital markets (foreign and domestic) is easily available
Ease of Doing Business
Banking regulation does not hinder competitiveness
Competition legislation in your country does not prevent unfair competition
Customs' authorities do not facilitate the efficient transit of goods
Financial institutions' transparency is not widely developed in your country
Easy to start company
Foreign investors are free to acquire control in domestic companies
Price controls affect pricing of products in most industries
Public sector contracts are sufficiently open to foreign bidders
Real corporate taxes are non distortionary
Real personal taxes are non distortionary
The exchange rate policy of your country hinders the competitiveness of enterprises
The legal framework is detrimental to your country's competitiveness
Protectionism in your country negatively affects the conduct of business in your country
Labour regulations hinder business activities
Subsidies impair economic development

Based on Kaufmann et al. (2007a)

4. The Worldwide Governance Indicators

- 1) All indicators from the same source are combined into a single indicator by taking a simple average. To illustrate, a simple average of the four indicators provided by the Economist Intelligence Unit for Regulatory Quality, namely “Unfair competitive practices”, “Price controls”, “Discriminatory tariffs” and “Excessive protections” is taken. Since all indicators from the same source have been combined into one indicator, there are now as many indicators as sources in the Regulatory Quality cluster.
- 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.
- 3) Each of the 15 sources for “Regulatory Quality” is classified as representative or non-representative depending on the distribution of the included countries across geographical regions and income. For “Regulatory Quality”, seven of the fifteen sources are classified as representative. 6 of them are expert assessments (5 commercial and the sixth based on commercial risk rating agency data) and one is a firm survey.
- 4) For the representative indicators in the Regulatory Quality cluster, the estimates of the unknown parameters, $\alpha(k)$, $\beta(k)$ and $\sigma_{\epsilon}^2(k)$, are obtained using Maximum Likelihood Estimation

(Equation 2)³⁹. 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 Regulatory Quality, $g(j)$, 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.

³⁹ 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.

4. The Worldwide Governance Indicators

- 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. Table 4.1 shows the weights for the WGIs in 2006.
- 10) The estimates of “Regulatory Quality” 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.

Table 4.1 Weights 2006 by Source and Worldwide Governance Indicator

Source	Voice and Accountability	Political Stability	Government Effectiveness	Regulatory Quality	Rule of Law	Control of Corruption
Commercial Business Information Providers						
Business Environment Risk Intelligence Business Risk Service (BRU)	..	0.09	0.09	..	0.07	0.01
Global Insight Global Risk Service (DRU)	..	0.12	0.04	0.03	0.03	0.03
Economist Intelligence Unit (EIU)	0.10	0.17	0.08	0.06	0.11	0.05
iJET Country Security Risk Ratings (UT)	..	0.09
Merchant International Group Gray Area Dynamics (MIG)	..	0.07	0.04	0.04	0.04	0.04
Political Risk Services International Country Risk Guide (PRS)	0.05	0.05	0.05	0.07	0.02	0.03
Business Environment Risk Intelligence Financial Ethics Index (QLM)	0.08	0.08
Global Insight Business Conditions and Risk Indicators (WMO)	0.05	0.17	0.15	0.17	0.09	0.07
Surveys of Firms or Households						
Afrobarometer (AFR)	0.03	..	0.10	..	0.01	0.02
World Bank Business Enterprise Environment Survey (BPS)	0.00	0.00	0.00	0.02
Transparency International Global Corruption Barometer (GCB)	0.02
World Economic Forum Global Competitiveness Report (GCS)	0.02	0.03	0.08	0.06	0.09	0.08
Gallup World Poll (GWP)	0.00	..	0.01	..	0.00	0.01
LatinoBarometro (LOB)	0.01	..	0.02	..	0.01	0.00
Political Economic Risk Consultancy Corruption in Asia Survey (PRC)	0.08
Institute for Management & Development World Comp. Yearbook (WCY)	0.01	0.05	0.03	0.07	0.07	0.11

4. The Worldwide Governance Indicators

Non-Governmental Organization Data Providers									
Bertelsmann Transformation Index (BTI)	0.12	..	0.07	0.09	0.03
Freedom House Countries at the Crossroads (CCR)	0.17	0.01	0.00
Global E-Governance Index (EGV)	0.01
Freedom House (FRH)	0.11	0.12	0.24
Global Integrity Index (GII)	0.09	0.02	0.00
Heritage foundation (HER)	0.03	0.05
International Research and Exchanges Board Media Sustainability Index (MSI)	0.12
International Budget Project Open Budget Index (OBI)	0.04
Reporters without Borders Press Freedom Index (RSF)	0.02
Public Sector Data Providers									
African Development Bank Country Policy and Institutional Assessments (ADB)	0.06	0.14	0.05	0.03
OECD Development Centre African Economic Outlook (AEO)	0.01	0.03
Asian Development Bank Country Policy and Institutional Assessment (ASD)	0.09	0.03	0.02	0.01
European Bank for Reconstruction and Development Transition Report (EBR)	0.10
Cingranelli Richards Human Rights Database and Political Terror Scale (HUM)	0.03	0.08	0.01
IFAD Rural Sector Performance Assessments (IFD)	0.01	..	0.03	0.03	0.01	0.02
World Bank Country and Policy Institutional Assessment (PIA)	0.04	0.07	0.05	0.05
US State Department Trafficking in People Report (TPR)	0.00

This panel reports the weights that would be used in the case of a hypothetical country appearing in all of the available underlying data sources for each indicator. Because of gaps in the country coverage of all of our data sources, no single country appears in all data sources. Nevertheless, the information reported in this panel is informative about the relative weights of the underlying indicators. Table based on Kaufmann *et al.* (2007a)

4.2.3 Characteristics of the WGIs

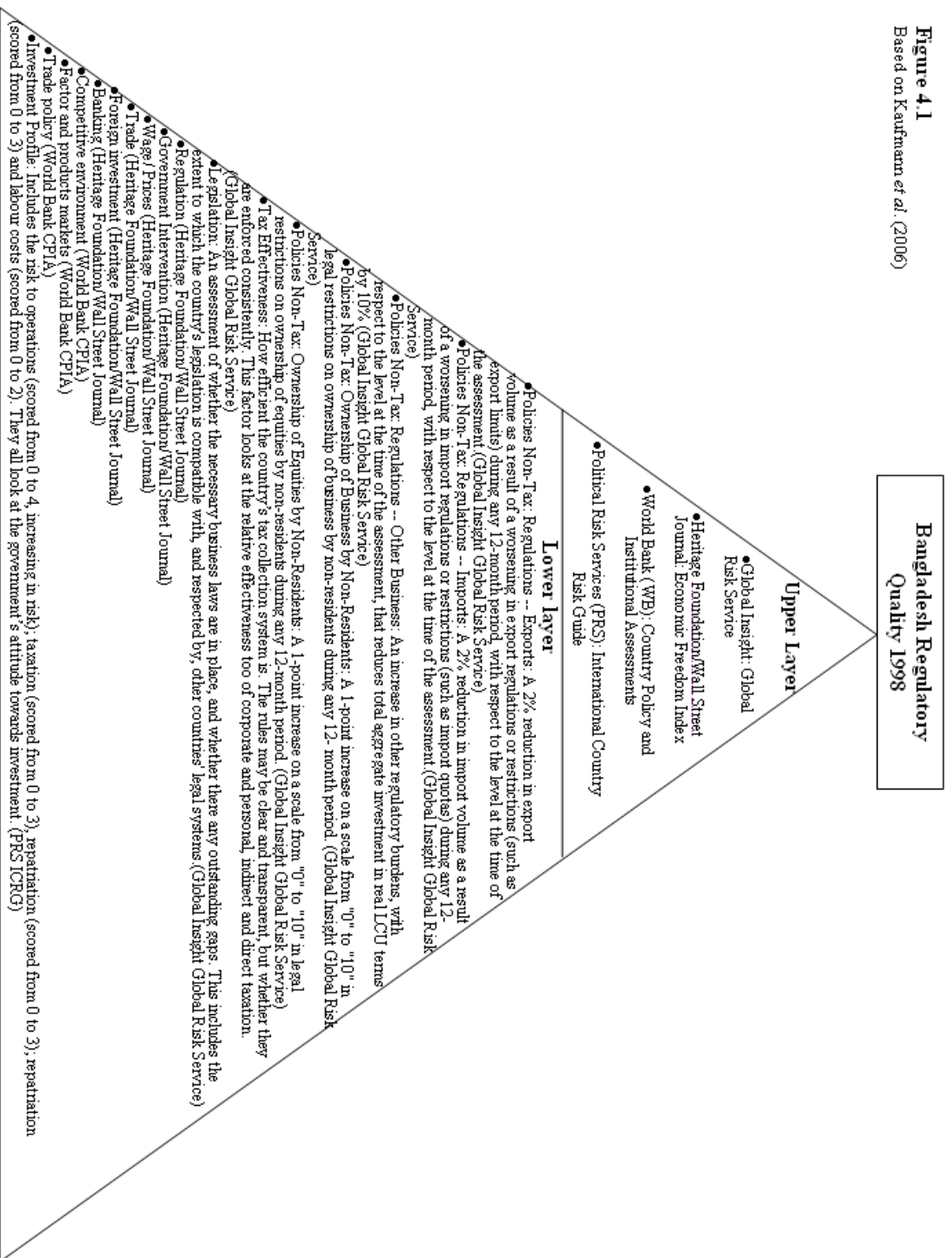
Three features are worth highlighting:

The first one is that the sources underlying a WGI differ across countries and over time. Figure 4.1 illustrates the change in sources over time. It shows the rating for Bangladesh on the WGI “Regulatory Quality” for 1998 and 2005⁴⁰. The upper layer shows the sources used to construct the rating, four in 1998 and 10 in 2005. The lower layer shows the indicators that underlie each source’s rating.

⁴⁰ I presented this example at a World Bank workshop in Bangladesh in November 2006.

4. The Worldwide Governance Indicators

Figure 4.1
Based on Kaufmann et al. (2006)



Bangladesh Regulatory Quality 2005

Upper Layer

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

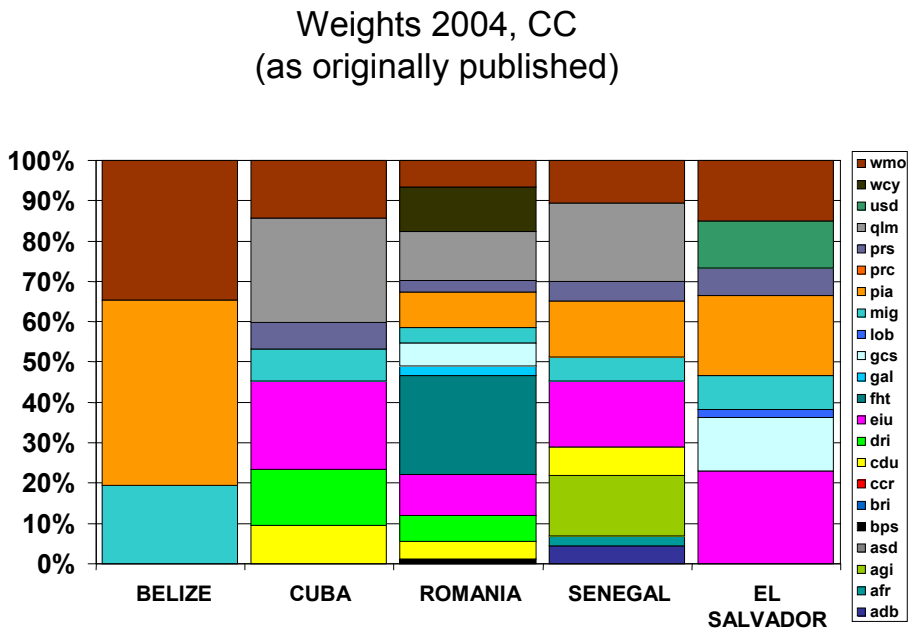
Lower layer

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

4. The Worldwide Governance Indicators

The second feature is that the weights assigned to each source differ across countries and over time, as illustrated in Figure 4.2. It shows the weights given to different sources for countries with similar scores on the 2004 “Control of Corruption” indicator. Larger countries with more sources tend to look more similar in terms of weights, although still not identical.

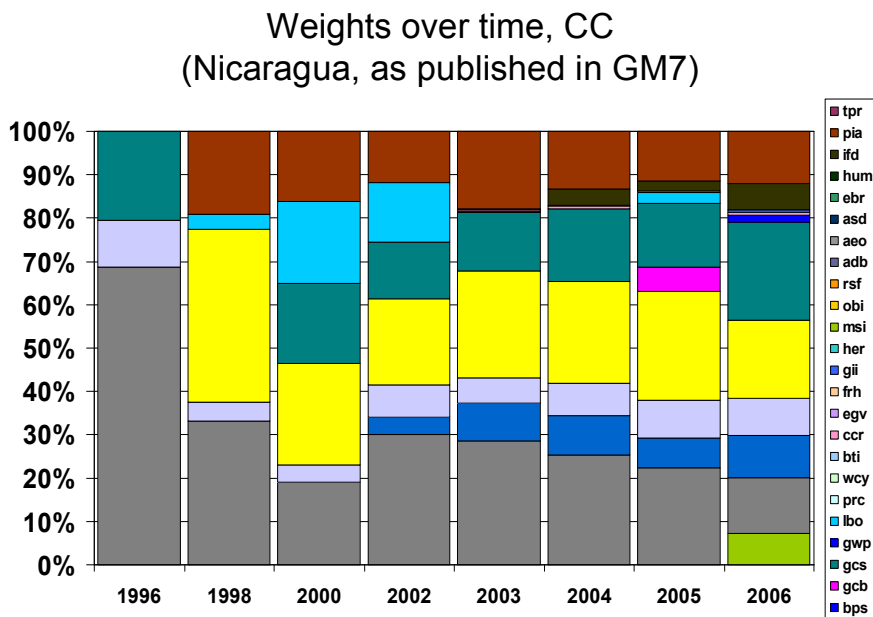
Figure 4.2



Source: Anderson (2007)

Figure 4.3 shows the weights for Nicaragua over time, again increasing them proportionally to make up for missing observations. In Nicaragua, 69% of the weight in 1996 went to EIU, while the same source got only 13% of the weight in 2006.

Figure 4.3



Source: Anderson (2007)

The reasons for this variation in weights are that *a)* the number and composition of sources differ across countries and over time and *b)* the degree of correlation between a source's ratings and the other sources' ratings changes over time.

The third feature of the WGIs is the publication of “margins of errors” along with the point estimates to represent the reliability of the rating. According to Kaufmann *et al.* (2005*b*), the standard error

4. The Worldwide Governance Indicators

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”. They construct 90 % confidence intervals on the basis of these margins of error around a country’s rating which, according to their estimate, has a 90 per cent probability of containing the country’s “true” (as opposed to estimated) score. Of course measurement errors are not unique to the WGI, but are present in all indicators.

Kaufmann *et al.* warn users not to compare countries among each other or over time where the confidence intervals overlap. To illustrate, users should not conclude that Bangladesh’s Regulatory Quality changed from 1998 to 2005, because the respective confidence intervals overlap (see Figure 4.1).

The margins of error and the corresponding confidence intervals decrease with i) an increase in the number of sources and ii) a higher correlation among the sources. To illustrate the first point, Kaufmann *et al.* (2003) explain that the 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 calculation of the standard errors is based on the assumptions that sources’ measurement errors are uncorrelated and that, therefore, a high correlation of sources reflects uniquely the high degree of accuracy of those sources, as opposed to less highly correlated sources. Hence, the more sources Kaufmann *et al.* use and the more these sources agree on countries’ ratings, the more reliable Kaufmann *et al.* believe their indicators to be. Also

important to note is the fact that an individual source's sample size is not taken into account in the computation of a source's error variance and accordingly not in the computation of its weight and the margins of error. To illustrate, a survey of more than 10 000 people is treated in the same way as a source that relies on the opinion of a few experts.

4.3 Concerns

The WGIs are a very useful first snapshot of experts' perceptions of a country's quality of governance. However, users often overestimate their explanatory power and use them for purposes they are not suited for. One major reason for the wide-spread misuse of the WGIs is that users are often unaware of one or all of the following five interrelated problems with the WGIs: i) violation of the key assumption of non-correlation of disturbances ii) lack of transparency of the ratings, iii) Weak conceptual foundation, iv) (Hidden) sample bias and contestable weighting method and v) lack of comparability over time and across countries. These points were raised in the book "Uses and Abuses of Governance Indicators" by Arndt and Oman (2006). Kaufmann *et al.* (2007b) refuted the criticism in their working paper "The Worldwide Governance Indicators Project: Answering the Critics". This section provides a reply to the arguments raised in "Answering the Critics", drawing on a working paper by Arndt, Knack and Oman (2007).

4.3.1 Correlation of disturbances

The key assumption behind the weighting of sources and the computation of the standard error is that the measurement error is uncorrelated across sources and across countries. 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.

This section discusses i) whether the assumption of non-correlated disturbances is realistic and ii) what the consequences of its violation are.

The likelihood of correlation of disturbances

There are several reasons why measurement errors tend to be correlated:

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

2. Experts who supply perception data for diverse sources are often informed of and influenced by perceptions and assessments from the same third parties
3. Perceptions used as inputs for the WGIs 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.

Examples⁴¹ include

- The WEF and IMD implement similar executive surveys, with samples selected by “partner institutes”. The WEF and IMD share many of the same partner institutes, so many of the same executives are likely to be included in both sets of surveys.

⁴¹ We note in passing that one of our examples in Arndt and Oman (2006) for the likely correlation of perception errors was that Kaufmann *et al.* use the Cingranelli and Richards Human Rights database and the Political Terror Scale as separate sources in the same indicators on the documentation of the WBI. We are glad to learn from Kaufmann *et al.* (2007b) that the documentation in the appendix in Governance Matters IV (Kaufmann *et al.*, 2005b, p. 103) and V (Kaufmann *et al.*, 2006, p. 84) where the two are mentioned as separate sources is misleading and that the two are actually treated as one source.

4. The Worldwide Governance Indicators

- The CPIA process takes into account numerous expert assessments and firm surveys, and ratings often are adjusted to be more consistent with rankings from those sources.
- The Economist Intelligence Unit, which is one of WGI'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 WGI "Control of Corruption" (Galtung and Sampford, 2005).
- Both the CPI and the ICRG corruption indicators showed corruption in Indonesia to be falling until 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).

"Interdependence does not even require, however, that sources directly check each other ratings. It can also result merely from sources relying on many of the same media reports or other qualitative sources of information about conditions in countries. If for example the international media cover corruption incidents more thoroughly for large (more newsworthy) countries than for small countries, the same unintentional bias against large countries could show up in the corruption ratings provided by different sources." (Knack, 2007, p. 266)

In earlier publications (1999b) Kaufmann *et al.* explained why correlation of perception errors is likely, and discussed to what extent the margins of errors would change. We are surprised therefore that Kaufmann *et al.* (2006, 2007a,b) more recently argue that, because they find “at most quite modest evidence of error correlation” (p. 19) and because there is no sufficient empirical evidence for the correlation of perception errors, our cautions about correlated errors do not hold.

Kaufmann *et al.* do not put any convincing arguments forward that discard the likelihood of serious correlation of perception errors:

1. The “empirical evidence” in Kaufmann *et al.* (2006) is based on the argument that, if perception errors are correlated, one would expect commercial risk rating agencies’ ratings to be highly correlated with each other, but less so with the ratings of surveys of firm. Kaufmann *et al.* report that the survey data are as highly correlated with expert ratings as the expert ratings are with each other, and conclude that the correlation of perception errors must not be serious. Their argument is based on a single selective and misleading case, the World Economic Forum’s executive opinion survey. No other survey source is correlated so highly with the expert ratings.
2. Moreover, perception errors can also be correlated across firm surveys and expert assessments. Experts could easily access

4. The Worldwide Governance Indicators

the World Economic Forum firm survey results, either directly through the WEF's publications or indirectly through TI's corruption index, which includes the most recent three WEF surveys. A better test would be to look at correlations between expert assessments and household surveys not included in TI's index, including TI's own Global Corruption Barometer.

3. The lack of sufficient evidence for the correlation of perception errors is a surprising argument for discarding our cautions. It certainly is not a reason to believe in the intuitively implausible assumption of uncorrelated perception errors.
4. Not only the correlation of sources' perception errors is problematic, but also the correlation of perception errors across countries. Kaufmann *et al.* do not discuss the latter in "Answering the critics".

Consequences of correlation of disturbances

The violation of the assumption that disturbances are not correlated has an impact on i) the consistency and unbiasedness of the estimates and ii) the size of the standard error. The assumption allowed Kaufmann *et al.* 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 variance⁴². 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. 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

⁴² 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).

4. The Worldwide Governance Indicators

among sources' errors from 0 to 0.5 and examining its effect on a sample of indicators and countries (Kaufmann *et al.*, 1999b). They report that the average standard error⁴³ *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 0.35⁴⁴.

For regression analysis, a higher measurement does not have important consequences for the validity of the findings. In fact, statistically, it can merely make it more difficult to find significant relationships in the data because of attenuation bias. In diagnosing governance problems and monitoring governance progress in a particular country, however, a higher than stated measurement error can have more serious consequences. Unfortunately it is impossible to know the "true" extent of measurement error of the WGI, because i)

⁴³ According to Kaufmann *et al.* (2005a), the standard error reflects the measurement error embodied in the point estimates for each country in a given year (see Box 4.2). Kaufmann *et al.* (2005a) refer interchangeably to "standard error" and "margins of error".

⁴⁴ 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.

the degree of correlation of disturbances is unknown and ii) the fact that the computation of the standard error is purely based on the number and correlation of sources is contestable, in particular if sources measure different aspects of governance.

Knack (2007, p.267) asserts that “the unknown but substantial degree of interdependence among many of the sources also obviates any claims regarding the “precision” of estimates. Other things equal, one can have more confidence in a rating based on 9 sources than on a rating for another country based on only 3 sources. It is also important however to identify the sources and to consider the likely degree of interdependence among them. Three sources consisting of a firm survey, a household survey and an expert assessment may provide a richer set of information than nine sources, if all nine are expert assessments. Iceland’s 2002 TI index is computed from six sources, which at first glance appears impressively diverse. However, none of them are truly independent: three of them are from WEF surveys for 2000, 2001 and 2002, and the other three are from IMD surveys for the same years. Although the partner institutes in Iceland are different for WEF and IMD, the likelihood of overlapping samples of top executives with international experience in a country so tiny must be very high. Iceland in TI is an extreme example of interdependence, but the problem in more moderate form is endemic to both TI and WBI. Claims of “being precise about imprecision” (Kaufmann, Kraay and Zoido-Lobaton, 2000) depend on independence of assessments, hence cannot be supported.”

4. The Worldwide Governance Indicators

Furthermore, 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. We would therefore be interested precisely in the non-correlated variation that is discarded by KKZ as measurement error. (See section on “sample bias and weighting method”). Conventionally, standard errors are based on the sample size, information which is available for few of the WGI sources and is discarded in the aggregation procedure.

While one may question the reliability of the assumptions reflected in the size of the standard errors and corresponding confidence intervals, it is equally important to emphasize that measurement errors are not unique to the KKZ indicators. On the contrary, KKZ are among the few producers of governance indicators who try to raise awareness among potential users of governance indicators of the serious measurement problems associated with such indicators.

4.3.2 Lack of transparency

Some donors have come to give high importance to the use of relatively untransparent indicators like the WGI or the CPI in, paradoxically, trying to give transparency to their aid-allocation

decisions. The reasons for the WGI's lack of transparency are threefold: i) high complexity, ii) lack of reproducibility, iii) lack of underlying theory of governance.

First of all, the sheer number and diversity of indicators, produced by others, that are incorporated into a single WGI make it very difficult to understand a country's rating, as illustrated in Figure 4.1 on page 111 for the WGI "Regulatory Quality" in Bangladesh 1998 and 2005. Understanding a change in a country's rating over time or a cross-country difference is complicated, because the composition and weights of primary sources vary across countries and over time (see Figure 4.2 and 4.3 on page 113). Knack (2007) explains that the WGI "Control of Corruption" uses 23 different combinations of sources for the Eastern European and Central Asian countries and that there are only four pair of countries whose ratings are based on a common set of sources. This difficulty to understand the meaning of a country's rating is all the greater because some of the underlying indicators used as inputs are themselves very broad and imprecise or lacking transparency. To illustrate, it is not clear what exact criteria the Economist Intelligence Unit (EIU) uses for rating countries on the indicator "Excessive Protections", one of the sources for Regulatory Quality in Bangladesh (see Figure 4.1). When are protections necessary and when are they excessive according to the EIU? Many sources unfortunately provide only headings for their indicators, with no definition or no document of how their assessments were made.

4. The Worldwide Governance Indicators

Secondly, the WGIs are not replicable. Many of the indicators underlying each source's ratings, the 2nd layer in the pyramid, are not disclosed to the public (see Figure 4.1 page 111). Access to many of the indicators of the 2nd layer would be very expensive. To illustrate, access to the source Global Risk Data from Global Insight, costs \$ 12 700 (e-mail inquiry 2004).

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

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

4.3.3 Weak conceptual foundation

It is important to remember that all rating systems are normative by their very nature. They imply a judgment, if they rate one country better than another country. They imply a judgment about what

government outcomes are desirable and what outcomes have priority over other outcomes. Given that governments' objectives are generally vague and necessarily self-contradictory (van de Walle, 2008), there is no single objective way of measuring the quality of governance.

One might, therefore, expect the WGIs to be based on a normative definition or concept of what constitutes good or bad governance which determines the countries' ratings. However, there is no such normative concept or theory underlying the WGI ratings. The authors define these ratings simply as "the traditions and institutions by which authority in a country is exercised". They further disaggregate the channels into different components, yet they do not explain how to distinguish between good or bad governance.

On the contrary, it's the compilation of all the ratings from different sources brought together in a single "WGI" which gives normative meaning to the concept of governance. But nobody, including the authors themselves, can define these norms. In addition, the norms differ from year to year and from country to country, because the sources for the WGIs, as for the CPI from Transparency International, differ across countries and over time (see Figures 4.1, 4.2 and 4.3).

This different mix of indicators over time and across countries without any conceptual explication also makes the interpretation of

4. The Worldwide Governance Indicators

the “margin of errors” unclear: “A statistic such as the confidence interval around a sample mean has a very specific interpretation: There exists a population of *something specific*. We estimate the mean of the population using observations on *that same specific thing*, but only for a subset of the population. Since we only have measurements for a sample, we arrive at a confidence interval for the true population mean for *that specific thing*. The WGIs have different mixes of indicators in different countries and in different years. We don’t have different observations on the same thing, we have different observations on different things. What the “margin of error” means in such case is unclear.” (Anderson, 2007)

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

Furthermore, rating countries on “true governance” implies that there is a best or ideally performing government (van de Walle, forthcoming). One may wonder about the nature of the opposite, “untrue governance,” and whether a single system of desired governance outcomes and priorities should be imposed on the developing world. Andrews (2008, p.2) analyses the content of the WGI “Government effectiveness” and argues that “the good governance picture of effective government is not only of limited use in development but also constitutes a threat, promoting isomorphism,

institutional dualism, and ‘flailing states’ and imposing an inappropriate model of government that “kicks away the ladder” today’s effective governments climbed to reach their current states”. Meisel (2004) explains why “Anglo-Saxon glasses” do not help to explain the rapid growth in the Asian economies and offers a new model of “Governance Cultures”, helpful for developing countries today. Chapter 5 will further elaborate on the current debate about “what governance is good for developing countries” and the challenges involved in measuring it.

4.3.4 Sample bias and weighting method

Sample bias

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

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

4. The Worldwide Governance Indicators

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

While the list of sources used to produce the KKZ indicators (see Box 4.1 on page 88) 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. To illustrate, Gallup’s World Poll that asks citizens about their exposure to crime gets zero weight for the WGI “Rule of Law”, whereas Global Insight Business Risk and Conditions, a commercial business information provider from Boston that measures the crime risk to businesses, gets the third highest weight. An example of a Global Insight indicator is “How much of a threat businesses face from crime such as kidnapping, extortion, street violence, burglary and so on. These problems can cause major inconvenience for foreign investors and require them to take expensive security precautions”.

Two reinforcing factors lie behind the low weights given to household surveys: first of all, there are fewer cross-country comparable household surveys and they provide data for fewer countries, because they are expensive and because there are few opportunities to exploit the data commercially. Kaufmann *et al.* use most cross-country comparable governance indicators available and commercial interests drive a major part of this supply of governance indicators.

Secondly, while the authors are not to responsible for the low number of household surveys available, they could have corrected for this selection bias in the aggregation process. Instead, they chose to use an aggregation technique⁴⁵ which gives less weight to perceptions that diverge from the dominant majority than to perceptions that conform to the majority view. Household surveys diverge from the majority of expert perceptions, and therefore, get low weights. This divergence is 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 surveys⁴⁶.

The reason for this divergence is not necessarily that either household surveys or expert assessments are wrong, but could also be

⁴⁵ This method should not be confused with factor analysis. See first part of this chapter for details

⁴⁶ Razafindrakoto and Roubaud (2006).

4. *The Worldwide Governance Indicators*

that they assess different aspects of governance. To come back to our previous example, the crime rate and protection (or abuse) by the police and judiciary are not necessarily at the same level for the poor parts of a city and the diplomatic and business parts. However, Kaufmann *et al.* assume that the discrepancy between Gallup's rating and the majority of expert ratings is due to its high measurement error and discard the possibility that it measures something different.

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 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 and Sampford, 2005).

This bias in favour of expert assessments (many 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, 2004a, p. 6). 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 regulations⁴⁷. The danger is that users of this indicator could *unwittingly* be influenced in

⁴⁷ 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".

4. The Worldwide Governance Indicators

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 spokesperson⁴⁸ 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".

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

⁴⁸ John D. Sullivan, Executive Director, Center for International Private Enterprise, Washington D.C.

collaboration between the agency and governments it was intending to help.

It is not problematic in general to produce an indicator that relies on the perceptions of experts and firms. It is even quite understandable to use these sources, because there are simply few cross-country comparable surveys of individuals available, and facts-based indicators are often no more objective than perceptions-based indicators (see Arndt and Oman, 2006). But it is misleading to label such an indicator differently. Kaufmann *et al.* market their indicators as “reliable measurements of governance”, whereas Transparency International (whose Corruption Perceptions Index relies on similar sources to the WGI “Control of Corruption”) correctly describes its indicators as “a snapshot of the views of business people and country analysts of corruption”.

Weighting method

Kaufmann *et al.* respond in “Answering the critics” to the criticism of their weighting method (raised in the sections on sample bias and correlation of perception errors) by using two alternative weighting schemes.” They report that the indexes produced with these alternative weighting systems are very highly correlated with their original WBI indicators, and conclude that criticisms of their weighting methods are of little importance.

4. The Worldwide Governance Indicators

However, neither of these two methods corrects for the fact that household surveys are in the minority, nor do they take into account that those sources that rate countries different from the majority of sources may actually contain very useful information about the quality of governance.

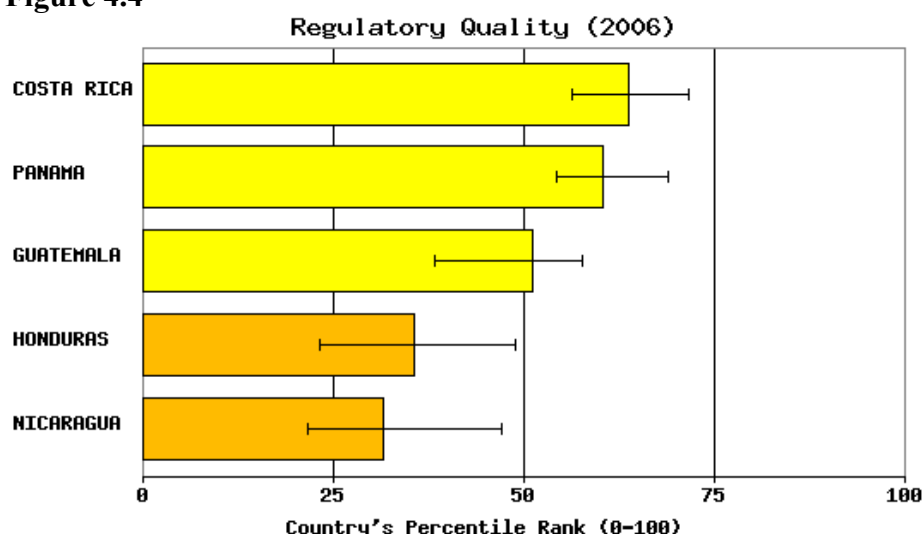
1. *(Un)equal weighting*: Kaufmann *et al.* use a simple average of all their sources and call this method “equal weighting”. Since there are many more expert assessments and firm surveys than surveys of individuals available, “equal weighting” is actually “unequal weighting” or “some are more equal than others”. The weight of household surveys is still negligible.
2. *Weighting according to groups*: Kaufmann *et al.* first take the average rating of four groups, i) cross-country surveys of firms and individuals, ii) commercial expert assessments, iii) expert assessments produced by NGOs and iv) expert assessments produced by multilateral organisations. They then simply average the indicators within each group and construct an aggregate indicator based on the four groups alone. Weighting according to groups is in general a very useful method to correct for an unbalanced sample of different types of sources. However, the fact that Kaufmann *et al.* put household surveys and firm surveys together in one group reduces the weight of household surveys to a not-even specified fraction. Household surveys should be taken as a separate group, because it is precisely household surveys whose results differ considerably

from the other sources. Furthermore, putting all surveys that often have a very high sample size (Executive Opinion Survey from the World Economic Forum relies for instance on more than 10,000 respondents) in one group and forming three separate groups for expert assessments, reduces the weight of surveys in general to only one fourth.

The claim that different weighting systems do not lead to substantially different results⁴⁹ therefore cannot be supported. Sources do measure different aspects of governance and rank countries individually differently; hence what sources receive more weight will determine the overall rating. For the purpose of illustration we will have a close look at "Regulatory Quality" 2006. Figure 4.4 shows the overall results of the rating. Table 4.2 shows how each of the ten sources rated the five Central American countries.

⁴⁹ "More systematically, as we discuss further below, the rankings provided by our aggregate indicators are quite robust to alternative weighting schemes" (Kaufmann *et al.*, 2007b, p. 13)

Figure 4.4



Source: Kaufmann D., A. Kraay, and M. Mastruzzi 2007: Governance Matters VI: Governance Indicators for 1996-2006

Note: The governance indicators presented here aggregate the views on the quality of governance provided by a large number of enterprise, citizen and expert survey respondents in industrial and developing countries. These data are gathered from a number of survey institutes, think tanks, non-governmental organizations, and international organizations. The aggregate indicators do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent. Countries' relative positions on these indicators are subject to indicated margins of error that should be taken into consideration when making comparisons across countries and over time.

Table 4.2

"Regulatory Quality" 2006 in Costa Rica, Guatemala, Honduras, Nicaragua and Panama

	Costa Rica	Guatemala	Honduras	Nicaragua	Panama
<u>Percentile Rank</u>	63.9	51.2	35.6	31.7	60.5
Source	Source rating (on a scale from 0 to 1)				
<u>Bertelsmann Transformation Index</u>	0.78	0.47	0.61	0.53	0.72
<u>Economist Intelligence Unit</u>	0.75	0.5	0.5	0.5	0.5
<u>Global Insight Business Conditions and Risk Indicators</u>	0.69	0.44	0.38	0.38	0.69
<u>Global Insight Global Risk Service</u>	0.95	0.87	0.89	0.72	0.95
<u>Heritage Foundation Index of Economic Freedom</u>	0.67	0.66	0.65	0.69	0.7
<u>IFAD Rural Sector Performance Assessments</u>	0.64	0.73	0.51	0.56	0.65
<u>Merchant International Group Gray Area Dynamics</u>	0.55	0.4	0.35	0.35	0.43
<u>Political Risk Services International Country Risk Guide</u>	0.64	0.82	0.55	0.73	0.82
<u>World Bank Country Policy and Institutional Assessments</u>	Not public	Not public	0.7	0.6	Not public
<u>World Economic Forum Global Competitiveness Report</u>	0.46	0.49	0.36	0.36	0.44

Scores from different individual indicators are not directly comparable with each other since the different data sources use different units and cover different sets of countries. The data from the individual indicators are further rescaled to make them comparable across data sources before constructing the aggregate governance indicators. See Aggregating Governance Indicators (Kaufmann et al., 1999b) for details.

Guatemala comes third out of the five Central American countries with the same sources. However, only two sources rank it third. Three sources rank it first (World Economic Forum Executive Opinion Survey, IFAD Rural Sector Performance Assessments,

4. The Worldwide Governance Indicators

Political Risk Service International Country Risk Guide,)), one ranks it second, two rank it fourth and one ranks it fifth. One of the sources that rank it third, Global Insights Business Conditions and Risk Indicators, has more weight than the three sources taken together that rank Guatemala first. If considerably more weight were given to those sources that rank Guatemala first, Guatemala's position on the rating would improve.

Panama comes second for "Regulatory Quality", after Costa Rica. However, out of the 9 sources we have data for, only four ranks Panama behind Costa Rica. Two give the same score to Costa Rica and Panama and three rank Panama higher than Costa Rica. Taken together, the four sources that rank Costa Rica higher than Panama have almost double the weight than the three sources that rank Panama higher than Costa Rica. Again, with different weights, Panama would improve its ranking.

The high divergence in the sources' ratings is reflected in the high margins of error: Guatemala's 90% confidence interval overlaps with the confidence interval of all other four countries. And Panama's confidence interval overlaps with Costa Rica and Guatemala (see Figure 4.4). Again, this reinforces the warning that rating countries on the WGIs, especially - but not only - where the confidence intervals overlap, is misleading.

4.3.5 Lack of comparability over time

The WGI cannot be used for direct comparisons of the quality of governance over time⁵⁰. An actual change in the level of governance is only one of four possible reasons for a change in a country's rating on the WGI.

A second possible reason is a change in the sources' perceptions of the quality of governance without an actual change in the quality of governance. For instance, in December 1996, a string of corruption cases hit Charles Haughey, who had been prime minister in Ireland from 1979 to 1992. The International Country Risk Guide, one of the sources for the WGIs, downgraded Ireland several times in the months that followed and during which more and more evidence of the scandals appeared. It is unlikely that the level of actual corruption had changed in Ireland, but the ICRG updated its rating based on newly available information.

⁵⁰ We note in passing that Kaufmann *et al.* misunderstood the reason for applying their rule of thumb - not to compare results over time or across countries with overlapping confidence intervals - in Arndt and Oman, p. 68. (see answer to "Critique 4", a critique we never made, in Kaufmann *et al.*, 2007, p.10): The aim of this section was to emphasize to users the importance of using Kaufmann *et al.*'s Rule of Thumb Rule which we very much appreciate. By showing what changes are significant at a 90 % significance level, we aimed at reinforcing Kaufmann *et al.*'s own cautions to only make comparisons over time if the margins of errors do not overlap. And we are somehow surprised that the authors go away from these claims. Given that the actual margins of error are likely to be much higher, so that more caution rather than less is appropriate, we find it inadequate to induce users to rely on a 75 percent significance level instead of a 95 % one (see p. 11 in "Answering the critics").

4. The Worldwide Governance Indicators

A third possible reason for a change is that other countries' governance ratings changed. Since Kaufmann *et al.* 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, changes in one country's rating change the ratings of other countries, other things equal.

A fourth possible reason is a change in the number and composition of sources from one year to the next. To illustrate, Guatemala's rating on Regulatory Quality worsened considerably from 1998 to 2005. The rating for 1998 was based on three sources and the rating for 2005 on eight sources, among them the three original sources from 1998. The rating of one of these three sources, the CPIA, is not publicly available for 1998. The other two sources that are available for both years actually rated the country better in 2005 than in 1998. The deterioration in Guatemala's rating from 1998 to 2005 was, hence, very likely driven by the additional sources for the 2005 rating (example based on data from Kaufmann *et al.*, 2006). Changes in the components of a composite indicator over time are not necessarily problematic for its comparability over time, if there are substantive reasons for the change (for instance, computers were added to the basket of goods used to calculate the Consumer Price Index, because consumption patterns change). However, changes in the WGI are not explained conceptually.

Contrary to their own earlier work (see Kaufmann *et al.*, 1999 a,b), Kaufmann *et al.* (2005b, 2006, 2007a,b) more recently claim

that their indicators actually do allow for direct comparisons of governance levels over time, if confidence intervals do not overlap. Their major argument is that their exercises based on a selection of sources turned up” little evidence of significant trends in world averages of governance” (2007b, p. 3). More precisely, Kaufmann *et al.* (2007b, p. 3.) claim that “the evidence from our individual sources that world averages of governance are not changing much is crucial, because it allows us to interpret the relative changes in country scores on our aggregate indicators, or groups of countries’ scores, as absolute changes.” We do not find this argument persuasive for the following reasons:

- The “evidence” that global averages do not change over time is based on the assumption that all underlying sources are comparable over time. However, many underlying sources used to construct the KKZ indicators do not allow for reliable comparisons of levels of governance over time: To illustrate, rating changes in expert-based systems do not always reflect a belief that actual conditions have changed, but often are intended to correct a previous year’s rating that in retrospect appears too high or too low. The documentation of the Heritage Foundation’s Property Rights Index, one of the sources for “Rule of Law”, frequently mentions the availability of new data as the core reason for changes in a country’s rating⁵¹ and World

⁵¹ See yearly editions of the Heritage Foundation’s “Index of Economic Freedom” book, from 1996 to 2008.

4. The Worldwide Governance Indicators

Bank staff involved in the construction of the CPIA affirmed that new ratings often correct last years' ratings.

- 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
- 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.

Kaufmann *et al.*'s also argue that i) the effect of changes in sources accounts for only a small portion of the variation in large changes and that differences in sources are therefore negligible and that ii) every source is a proxy for the same "true governance" and that hence the use of different sources for every year does not limit the comparability over time.

We do not support these claims. Since especially the changes in governance ratings over time can have far-reaching consequences for resource allocations to developing countries and for the reputation of a country's government, it is important not to downplay the impact of a change in sources, but to encourage those using the indicators to always check what factor(s) drive(s) the change in a country's rating. While changes in sources may on average not be the main driver of changes in the ratings, they still contribute to many changes over time

and in some case are even the sole drivers of changes over time. To illustrate, the WBI indexes for 2002 and 2004 are based on the same set of sources for only 4 of the 27 countries in the ECA region: Hungary, Slovak Republic, Slovenia and Tajikistan. Furthermore, as will be shown in the next section, sources measures different things and different compositions lead to different results. Hence fundamental changes in sources can also fundamentally change the meaning of an indicator (which is always implicit and hence the change is not visible to the user of the indicators).

An illustration of the sensitivity of the ratings to the inclusion of new sources is the case of the WGI "Regulatory Quality" in El Salvador and Bolivia from 1998 to 2005, as highlighted in Figure 4.5. While Kaufmann *et al.* found in their 2006 publication "Governance Matters V" a statistically significant deterioration of "Regulatory Quality" from 1998 to 2005 (see Figure 4.5a) in both countries, they do not find one in the recent update of their work (Kaufmann *et al.*, 2007a, "Governance Matters VI", see Figure 4.5b) for the same period. The difference in the two graphs is mainly due to a difference in the rating for the year 1998, due to a change in sources from "Governance Matters V to VI" (see Table 4.3). In Figure 4.5a), the rating for 1998 is based on three sources for El Salvador and five sources for Bolivia, while in Figure 4.5b) it is based on six sources for El Salvador and seven sources for Bolivia. The additional sources change the percentile rank for 1998 dramatically, from 94.1 to 78 in the case of El Salvador, and from 78.8 to 60 in the case of Bolivia. Due to the change in rating and confidence interval for 1998, the over time

4. *The Worldwide Governance Indicators*

change from 1998 to 2005 is smaller and not statistically significant anymore.

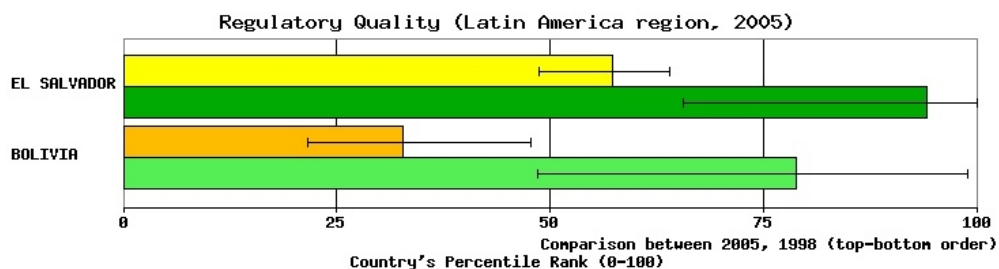
Transparency International, whose CPI suffers from similar problems, clearly warns its users that “Year-to-year changes in a country's score can either result from a changed perception of a country's performance or from a change in the CPI's sample and methodology. The only reliable way to compare a country's score over time is to go back to individual survey sources, each of which can reflect a change in assessment”⁵². This also applies to the World Bank Institute's WGI. Transparency International acknowledges that its CPI at best reflects changes in perceptions of the quality of governance over time. Kaufmann *et al.* in contrast choose to make much stronger claims about measuring changes in governance.

52

http://www.transparency.org/policy_research/surveys_indices/cpi/2006/faq#using1

Figure 4.5 a)

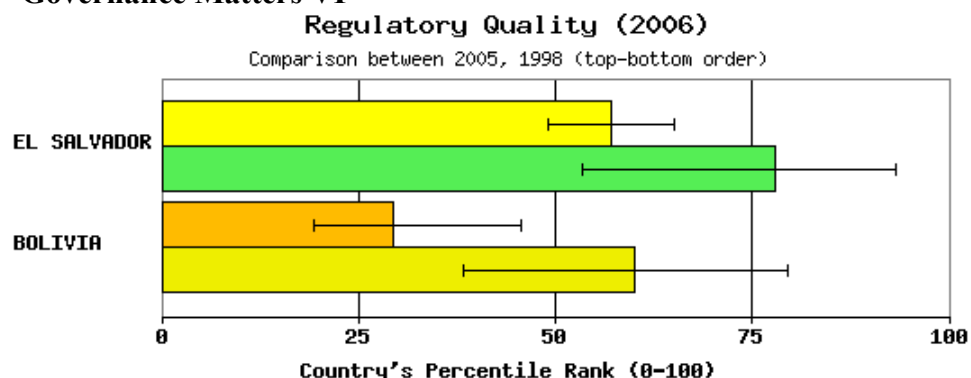
**"Regulatory Quality" 1998 and 2005 in El Salvador and Bolivia,
"Governance Matters V"**



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

Figure 4.5 b)

**"Regulatory Quality" 1998 and 2005 in El Salvador and Bolivia,
"Governance Matters VI"**



Source: Kaufmann D., A. Kraay, and M. Mastruzzi 2007: Governance Matters VI: Governance Indicators for 1996-2006

Note: The governance indicators presented here aggregate the views on the quality of governance provided by a large number of enterprise, citizen and expert survey respondents in industrial and developing countries. These data are gathered from a number of survey institutes, think tanks, non-governmental organizations, and international organizations. The aggregate indicators do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent. Countries' relative positions on these indicators are subject to indicated margins of error that should be taken into consideration when making comparisons across countries and over time.

Source: based on data in Kaufmann *et al.* (2007a)

Table 4.3 a)
Sources for "Regulatory Quality" 1998 El Salvador and Bolivia in
"Governance Matters V" Data for figure 4.5a)

El Salvador "Regulatory Quality" 1998 (score of 1.31 on a scale from -2.5 to +2.5)				
	Source	Publication	Type	Source's rating (scale from 0 to 1)
1	Heritage Foundation/Wall Street Journal	Economic Freedom Index	Poll	+0.75
2	World Bank	Country Policy and Institutional Assessments	Poll	Not public
3	Political Risk Services	International Country Risk Guide	Poll	+0.83
Bolivia "Regulatory Quality" 1998 (score of 0.8)				
1	Global Insight	Global Risk Service	Poll	+0.88
2	Heritage Foundation/Wall Street Journal	Economic Freedom Index	Poll	+0.67
3	World Bank	Country Policy and Institutional Assessments	Poll	Not public
4	Political Risk Services	International Country Risk Guide	Poll	+0.83
5	World Bank	World Business Environment Survey	Survey	+0.62

Table 4.3 b)
Sources for "Regulatory Quality" 1998 El Salvador and Bolivia in
"Governance Matters VI" Data for figure 4.5b)

El Salvador "Regulatory Quality" 1998 (score of 0.74 on a scale from -2.5 to +2.5)				
	Source	Publication	Type	Source's rating (scale from 0 to 1)
1	Economist Intelligence Unit	Country Risk Service	Experts	0.55
2	Global Insight	Business Conditions and Risk Indicators	Experts	0.63
3	Heritage Foundation	Index of Economic Freedom	Experts	0.8
4	Political Risk Services	International Country Risk Guide	Experts	0.83
5	World Bank	Country Policy and Institutional Assessments	Experts	Not public
6	World Economic Forum	Global Competitiveness Report	Survey	0.45
Bolivia "Regulatory Quality" 1998 (score of 0.35)				
1	Economist Intelligence Unit	Country Risk Service	Experts	0.6
2	Global Insight	Business Conditions and Risk Indicators	Experts	0.63
3	Global Insight	Global Risk Service	Experts	0.88
4	Heritage Foundation	Index of Economic Freedom	Experts	0.7
5	Political Risk Services	International Country Risk Guide	Experts	0.83
6	World Bank	Country Policy and Institutional Assessments	Experts	Not public
7	World Economic Forum	Global Competitiveness Report	Survey	0.34

4.4 Conclusion

While the WGIs are a very useful first snapshot of experts' perceptions of a country's quality of governance, users often fail to take into account their limitations with respect to robust and meaningful cross-country and over time comparisons. The high complexity of the indicators hides the normative assumptions behind them – which differ across countries and change over time due to different compositions of the indicators.

The debate about those implicit assumptions and about more “actionable” indicators would touch the very heart of the discipline of development: What governance is needed for developing countries? The fact that there is no easy answer to this question might go far in explaining the currently low number of more transparent and actionable indicators. The following chapters build on the first four chapters: While chapter 5 looks deeper into the politics behind governance ratings and discusses alternatives and future trends, chapter 6 relies on cutting-edge econometric techniques and a unique panel dataset to test the common belief that a good “Rule of Law” drives economic growth.

5 The politics of governance ratings⁵³

⁵³ This article has been published in a special issue on international ratings in the International Public Management Journal (Arndt, 2008).

The last 15 years have seen a veritable explosion of interest in the quality of “governance” in the developing world. This very growth has driven an equally significant increase in the supply and use of quantitative governance indicators. The World Bank plays a leading role, both in the use and production of cross-country comparable governance indicators, followed by other international organizations. Though there are hundreds of governance indicators, a few dominate the market. World Bank indicators such as the Worldwide Governance Indicators (WGIs) and the Doing Business indicators are very popular outside the World Bank among researchers for regression analysis, bilateral donors for aid-allocation decisions, international investors who wish to analyse country risk for their investment decisions, and media that report on countries’ ratings with respect to their neighbours or previous years (Arndt and Oman, 2006). Other World Bank indicators such as the Business Environment and Enterprise Performance Surveys (BEEPs), the Country and Policy Institutional Assessments (CPIA) and numerous new, more actionable indicators are less frequently used outside the World Bank.

Previous works criticized the conceptual and technical limitations of the most popular indicators, the WGIs, and their misuse (Arndt and Oman, 2006; Charles *et al.*, 2008; Knack, 2007; Kurtz and Schrank, 2007; Thomas, 2007; van de Walle, 2005; see Kaufmann, Kraay, and Masstruzzi, 2007b for a reply) and the Doing Business Indicators (Arrunada, 2007; Berg and Cazes, 2007; see Djankov, 2008 for a reply). Hood, Dixon, and Beeston (2008) explain the recent growth in international rankings of Public Service Performance and

rate the rankings.

This paper attempts to contribute to the existing literature in the following ways. First, it draws on in-depth interviews of staff at the World Bank and other international organizations to understand why the World Bank produces cross-country comparable indicators and whether its staff believes that those indicators satisfy the different purposes for which they were created. It brings to light the current controversies within the World Bank about existing indicators and future work on governance indicators. The outcome of this debate will heavily impact World Bank policies in developing countries, not least because indicators determine aid-allocation and are expected to play a crucial part in setting the reform agenda in developing countries and assessing the success of reforms undertaken. Second, the paper attempts to explain why so many external users rely on the Worldwide Governance Indicators, despite the limitations of these indicators, the reluctance within the World Bank to use them and the availability of more meaningful indicators.

Many of the findings in this paper are based on twelve interviews with World Bank staff, undertaken in 2007 and 2008 and interviews with other development experts, undertaken between 2005 and 2008 (ten interviews with international banks, six interviews with bilateral donors, five interviews with staff of other international organizations, eight interviews with academics). Information was obtained from either or both questionnaires and (unstructured) interviews.

The paper is organised around three sections: the first section introduces the most popular as well as less well-known cross-country comparable indicators produced by the World Bank. The second section attempts to explain the motivations within the World Bank to produce governance indicators and whether they satisfy the expectations. The third section attempts to explain why use of the WGIs is so widespread, in particular outside the World Bank, despite the availability of alternatives. The paper ends with a conclusion and recommendations for the short and long run to consumers and producers of governance indicators.

5.1 World Bank governance indicators

World Bank President James Wolfensohn decided in 1996 to overturn the Bank's longstanding policy that it could not explicitly recognize or seek to address the acute problems of corruption in many of its borrowing countries, because local politics were outside the Bank's official mandate, and to give those problems a high priority. While World Bank lending to promote economic reforms fell by 14 per cent annually between 2000 and 2004, its lending to improve governance rose by 11 per cent annually during that period, so that by 2004, 25 per cent of its lending was committed to law and public administration in borrowing countries (World Bank, 2004).

The World Bank plays a front-runner role in the quantification of countries' quality of governance and in particular the WGIs (Arndt

5. The politics of governance ratings

and Oman, 2006) and the Doing Business indicators are very popular externally: investors increasingly want back-up to traditional, macroeconomic-based, country-risk indicators that have failed to predict costly financial crises. They increasingly base their investment- location decisions on their perceptions of governance and these perceptions are increasingly shaped by governance indicators. To illustrate, the use of the Worldwide Governance Indicators in country risk ratings is common (Arndt and Oman, 2006). Similarly, donor agencies increasingly rely on governance indicators for their aid-allocation decisions. The US Millennium Challenge Corporation, for instance, allocates billions of dollars of aid on the basis of governance indicators. It wants to identify and reward developing countries that are improving their governance and aims at increasing the transparency and consistency of its criteria for determining recipients' aid eligibility. The WGIs play a predominant role in these ratings. Other indicators used by the MCC include the Doing Business indicators. Development analysts and academics widely use for their quantitative studies the WGIs (Iqbal and Shah, 2008) and the Doing Business indicators, though the latter to a lesser extent because they only concern one specific aspect of governance. Journalists frequently report countries' ratings with respect to their neighbours and over time on both the WGIs and the Doing Business indicators.⁵⁴

⁵⁴ See for instance the section on media coverage on the official WGI website, http://worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0,,contenMDK:21407860_pagePK:64168445_piPK:64168309_theSitePK:1740530,00.html (accessed May 2008).

5.1.1 The Worldwide Governance Indicators

Huther and Shah (1996, 1998) were the first to develop an aggregated indicator of the quality of governance at the World Bank's Independent Evaluation Group for a sample of eighty countries, for one period in time. The aim of this work was to be able to judge the success of the Bank's new anti-corruption and governance strategy. Surveys to assess citizens' evaluations of governance outcomes were planned as a continuation of the work, but the work was not completed due to some, not further specified, "intervening factors." (Iqbal and Shah, 2008, p.5)

In 1999, Dani Kaufmann, Aart Kraay, and Pablo Zoido-Lobaton⁵⁵ published for the first time the World Governance Indicators (WGIs) at the World Bank Institute (Kaufmann *et al.*, 1999). The six WGIs, namely "Voice and Accountability," "Political Stability," "Government Effectiveness," "Regulatory Quality," "Rule of Law," and "Control of Corruption" summarize data from more than 30 expert assessments and household and firm surveys (including indicators from the World Bank's Country and Policy Institutional Assessments), and are available from 1996 to today for more than 200 countries. They are publicly available and easily accessible on the web.

⁵⁵ Currently, the WGIs are constructed by Dani Kaufmann, Aart Kraay and Massimo Mastruzzi.

5. *The politics of governance ratings*

According to one of the authors,⁵⁶ the WGIs were created in response to four interrelated concerns. One was the apparent lack of robustness of cross-country comparisons using different individual data sources, especially when the different sources led to different conclusions. Second was concern about how to interpret cross-country differences and their statistical and practical significance. Third was concern that it is difficult to compare results from regional surveys with broader cross-country surveys. A fourth concern was to find a way to produce useful overarching, integrative or summary indicators, given the large and growing diversity of individual sources cropping up in different pieces of research and in policy debates.

Publications and the website of the WGIs contain the following disclaimer: “The data and research reported here do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent. The WGI are not used by the World Bank Group to allocate resources or for any other official purpose.”

5.1.2 The Doing Business Indicators

The Doing Business Indicators were published for the first time in 2004 and are publicly available from 2003 onwards on a yearly basis for more than 100 countries (178 countries in 2008) by Simeon Djankov and his team at the International Finance Corporation of the World Bank (Djankov *et al.*, 2005). The overall rating on the ease of

⁵⁶ We thank Aart Kraay for this information.

doing business can be disaggregated into individual components of “starting a business,” “dealing with licenses,” “hiring and firing,” “registering property,” “getting credit,” protecting investors, paying taxes, trading across borders, enforcing contracts, and closing a business. Ratings are based on lawyers’ assessments of the *de jure* situation in a country. An indicator of “hiring and firing” therefore is informative about perceptions of how easy it is to hire and fire an employee according to the legal procedures that have to be followed and not of how long it may take in reality. Countries can challenge and seek corrections of the ratings. The methodology and criteria on which countries are ranked are openly disclosed on the web.

5.1.3 Other World Bank indicators

Many other cross-country comparable World Bank indicators exist, with varying data coverage. The World Bank Institute webpage lists twelve further governance datasets and empirical tools which are accessible⁵⁷. One dataset is the Business Environment and Enterprise Performance Survey (BEEPs). BEEPs is a joint initiative with the European Bank for Reconstruction and Development. It surveyed 20,000 firms in 3 rounds (1999, 2002, 2005) in 26 transition countries, and 6 European comparators in 2005 (Ireland, Germany, Greece, Portugal, Spain, Turkey) as well as Korea and Vietnam. BEEPs contains variables on unofficial payments and corruption, crime,

⁵⁷ <http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0>

” contentMDK:20749087_isCURL:Y_menuPK:1928447_pagePK:64168445_piPK:64168309_theSitePK:1740530,00.html (accessed May 2008).

5. The politics of governance ratings

regulations and red tape, customs and taxes, labour issues, firm financing, legal and judicial issues, and infrastructure. All data are available on the internet plus full reporting of the sample characteristics.

The most prominent dataset besides the WGIs and the Doing Business indicators is the annually produced Country Policy and Institutions Assessments (CPIA) which assesses the quality of Bank borrowing countries' policy and institutional frameworks for fostering poverty reduction, sustainable growth, and effective use of development assistance with sixteen indicators. Country ratings are based on World Bank staff judgments. While the focus of the CPIA in the 1970s and 1980s was largely on macroeconomic indicators, there are since the end of the 1990s several governance indicators in the CPIA, which play a crucial role in the World Bank International Development Association's (IDA) aid-allocation decisions. The cluster "Public Sector Management and Institutions" contains indicators on "Property Rights and Rule-based Governance," "Quality of Budgetary and Financial Management," "Efficiency of Revenue Mobilization," "Quality of Public Administration," and "Transparency, Accountability, and Corruption in the Public Sector." The CPIA is produced for IDA performance based-allocation, but is also used "to inform the Bank's Country Assistance Strategy program and country policy dialogue; to assist in the assessment by the Bank's Quality Assurance Group of the degree of portfolio risk; to help identify countries for extra attention on fiduciary standards and governance, to provide background to the Bank's Independent Evaluation Group's Country Assistance Evaluations, to help assess a

given country's debt sustainability, and to offer input to research on the determinants of growth and poverty reduction.”⁵⁸ 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 finalized and governments are consulted before the final rating is made public. The CPIA is mostly used internally, because data on each of the sixteen indicators for the most recent years was disclosed only recently for the IDA countries.

5.1.4 Move towards “actionable” indicators

The World Bank is currently involved in several projects that aim at developing more transparent, meaningful, and “actionable” indicators, i.e., indicators that are informative about what reforms countries could undertake to improve their ratings and that track progress over time. One prominent example is the Public Expenditure and Financial Accountability (PEFA) initiative, which is a partnership between the World Bank and other bi- and multilateral donors. Initiated in 2001, it measures critical dimensions of open and orderly public financial management systems in 28 indicators that capture six dimensions of public financial management. So far 67 assessments

⁵⁸ [http://web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/IDA/0, contentMDK:20941073_pagePK:51236175_piPK:437394_theSitePK:73154,00.html](http://web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/IDA/0,contentMDK:20941073_pagePK:51236175_piPK:437394_theSitePK:73154,00.html) (accessed May 2008).

5. The politics of governance ratings

have been completed, of which 26 are publicly available (PEFA, 2008).

Another example is the Global Integrity Index which is supported by the World Bank. It assesses the existence, effectiveness, and citizen access to key national-level anti-corruption mechanisms used to hold governments accountable. It is generated by aggregating more than 300 Integrity Indicators systematically gathered for each country covered. For the Global Integrity Index 2007, those indicators comprised more than 15,000 peer-reviewed questions and answers scored by in-country experts in the summer of 2007. Several rounds of review are conducted at the international level to ensure that cross-country comparisons are valid. In addition, all assessments are reviewed by a country-specific, double-blind peer review panel comprising additional local and international subject matter experts. All the underlying data and criteria used to rank the 55 countries are freely accessible to the public.

For the Latin American region, the World Bank plans to develop comparative indicators that benchmark the structure and performance of government bureaucracies in the field of audit, public financial management, intergovernmental finance, and justice. According to the project proposal, the project is designed to help Latin American countries benefit from the experience of their peers and OECD countries and will be complemented by demand-driven technical peer assistance.

Other indicators under development aim to assess “conflict of interest rules,” “asset monitoring,” “immunities,” and “freedom of information.”

While the World Bank produces many of the most popular and well-known indicators, there are many more indicators available from other producers. The Inter American Development Bank (IADB, 2007) lists more than 400 freely accessible indicators in its database.

5.2 Why does the World Bank produce cross-country comparable governance indicators? Do their indicators satisfy the purposes they are produced for?

The World Bank and other international organizations produce cross-country comparable composite governance indicators for a range of purposes. Governance ratings could help identify areas of reform and measure the success of governance reforms, could create incentives in developing countries to improve governance and put governance on the agenda, could contribute to country dialogue, could enhance the transparency and objectivity of aid-allocation decisions and could enhance quantitative governance research. Institutional reasons may also play a role. Whether the most-widely used governance indicators fulfil these expectations is the subject of an ongoing debate, both within and outside the World Bank. There is no consensus within the World Bank on which indicators are most useful. Rather, their respective strengths and weaknesses are debated in many seminars at the World Bank.

5.2.1 Help identify areas of reform and measure the success of governance reform

The externally most popular WGIs are generally not recognized as being useful to help identify areas of reform and measure the success of governance reforms. The authors of the indicators “caution users that the aggregate indicators can in some circumstances be rather a blunt tool for policy advice at the country level” (Kaufmann, Kraay, and Mastruzzi, 2007a, 23). Some other World Bank colleagues (Iqbal and Shah, 2006, 2008; Knack, 2007) and a series of academic works from outside the bank (e.g., Arndt and Oman, 2006; Kurtz and Schrank, 2007; Pollitt, forthcoming; Thomas, 2007; van de Walle, 2005, 2006) are more outspoken about the lack of transparency, clarity, conceptual validity, reliability, and comparability across countries and over time (see Kaufmann *et al.*, 2007b, for some comments on some of the critics). Thomas (2007) for instance finds no evidence for the construct validity of the WGIs. She concludes that the “indicators stand as an elaborate and unsupported hypothesis” and that “accordingly, reliance on the indicators is premature” (p. 2).

The changing compositions of the WGIs due to changes in the number of sources and weights across countries and over time makes it difficult for developing countries to understand what they should do to change their rating. World Bank staff member James Anderson illustrates that while 69% of the weight in 1996 went to the Economist Intelligence Unit for Nicaragua’s rating on the “Control of

Corruption,” the same source got only 13% weight by 2006. Similarly, while Belize would have benefited from a better CPIA rating which counts for 46% of its overall rating on the “Control of Corruption” in 2004, the influence of the CPIA rating is zero for Cuba, where this indicator is not included at all (Anderson, 2007). World Bank researcher Steve Knack (2007) highlights a range of additional reasons that limit the comparability over time of the WGIs.

Since sources measure different aspects of governance, it does matter what sources provide data for a country and how much weight is assigned to them. According to a World Bank staff member interviewed, it is unlikely that even a very successful reform in an area like tax administration will show up in the rating, given that an indicator such as “Government Effectiveness” comprises such diverse indicators as quality of infrastructure and public health and education services as well as indicators of red tape with changing weights assigned to them.

A further criticism within and outside the World Bank is that not all disaggregated data is publicly available. Iqbal and Shah (2008) estimate total costs of buying access to the 10 commercially provided sources of the WGIs at \$90,000. Furthermore, even if data for all underlying indicators were available, it is unclear what indicators such as “excessive protections” from the Economist Intelligence Unit means. When are protections excessive, and when are they deemed necessary? Since the indicators are not constructed on the basis of a

5. *The politics of governance ratings*

theory (Thomas, 2007), the normative content is implicit and depends entirely on the sources (Arndt and Oman, 2006).

Firm surveys such as BEEPs were cited as extremely useful for inducing changes within countries and measuring progress over time. In one country, the driving leader of change in the government continuously referred to the BEEPs to justify its reforms. On the other hand, World Bank staff cautioned that those surveys—while being very useful to gain an understanding of firms’ perceptions of the quality of governance in a country—have their limits and should not be used on their own for far-reaching reform decisions.

The Doing Business indicators are straightforward in their reform advice. They ask for a simplification of the procedures for setting up and formalizing firms and a deregulation of labour laws to make it easier to fire and hire employees. However, whether such reform advice is helpful has been contested. Arrunada (2007) argues that such reforms have disregarded the role business registers play as a source of information for judges, government, departments and other firms. Such information is deemed essential by the author to reduce transaction costs in future dealings with both public and private economic agents. Berg and Cazes (2007) argue that the indicators on employment are not satisfactory for formulating policy recommendations and, within the World Bank, Kaufmann (2004) draws attention to the large gap between results on the *de jure* Doing Business indicators and other, more *de facto* indicators.

There is disagreement within the World Bank with respect to the usefulness of the CPIA for identifying areas of reform and monitor progress over time. While some World Bank staff interviewed are convinced of this instrument, others expressed clearly their opposition (see also Herman, 2004; World Bank, 2001).

Much hope is put into the development of the new, more robust, and actionable indicators. The recent World Bank Governance and Anti-corruption Report (World Bank 2007, 9) takes a somewhat cautious line concerning the WGIs when it notes that they “exist alongside many other products (e.g., enterprise surveys) to provide comparative information for policymakers; they are not official ratings by the WBG, and they have no formal role in any WBG operational decision-making.” It notes at the same time that the “Development Committee asked the Bank to further develop and use disaggregated and actionable indicators. Disaggregated and actionable indicators can serve two purposes—to inform the CPIA and to help track progress in specific reforms implemented by governments.”

World Bank staff interviewed argued that one of the reasons for the slow progress towards “actionable” indicators is that it is not clear what is “action worthy.” While everyone may agree that a well-functioning and stable law system with guaranteed property rights, a low crime rate, a low level of corruption and a well-functioning bureaucracy are characteristics of developed countries, there is no consensus on which reforms should be the priority of governments in developing countries, given that the government has limited resources

5. The politics of governance ratings

at their disposal, both in terms of time and money. Researchers and practitioners worry that the current governance agenda is overloaded (Grindle, 2007) and that it can divert attention away from other, more urgent problems. Rodrik (2008) argues for “second-best institutions” instead of the current overambitious governance agenda. The former World Bank staff member Matthew Andrews, now at Harvard University, argues that “the good governance picture of effective government is not only of limited use in development but also constitutes a threat, promoting isomorphism, institutional dualism, and ‘flailing states,’ and imposing an inappropriate model of government that ‘kicks away the ladder’ today’s effective governments climbed to reach their current states.” (Andrews, 2008).

A debate about these issues has started in the World Bank. This is reflected in a recent seminar at the World Bank, and a volume produced by the World Bank’s International Bank for Reconstruction and Development (World Bank, 2008b) where leading scholars in the field, Douglass North, Daron Acemoglu, Francis Fukuyama and Dani Rodrik were invited to reflect upon the current stage of research on “Governance, growth, and development decision-making” and its implications for World Bank work on governance. To illustrate, while Acemoglu emphasizes that institutions matter for a country’s development, he also warns that there is no general recipe for improving institutions and that exporting good economic institutions is as hopeless as exporting democracy. Rodrik points to the importance of distinguishing between governance as an end in itself and governance as a means to achieve other desirable objectives and

the tradeoffs that can arise between the two. He criticizes the institutional reforms promoted by the World Bank and other multilateral organizations as being biased towards a best-practice model and underlines the importance of taking local constraints and opportunities into account.

5.2.2 Incentives for developing countries to improve their governance and agenda-setting

Governance ratings could help to direct attention within the World Bank and in developing countries towards governance. Governance ratings can also create incentives in developing countries to improve their governance in several ways: First, by “naming and shaming” through the media those countries with bad governance, it could give governments an incentive to pursue reforms and start a dialogue to improve the government’s and the country’s reputation. A country’s reputation plays a role for investment decisions and a bad performance of a country vis-à-vis its neighbours could lower the government’s popularity within the country. Countries may therefore engage in a “horse-race,” i.e., compete with each other, to obtain a better rating. Second, aid-allocation conditional on the quality of governance and/or its improvement could induce governments seeking aid to reform to improve their rating.

Transparency International’s Corruption Perceptions Index (CPI) was successful in drawing attention to the problem of corruption and in putting it up front on the agenda (TI 2004). According to some

5. The politics of governance ratings

staff members interviewed for this study, the WGIs were very helpful in lobbying within the World Bank for increasing its work on governance. Local media also widely report on countries' ratings on the Worldwide Governance Indicators and the Doing Business indicators and put local governments under pressure. A staff member of the Millennium Challenge Corporate reported that governments started a debate on the basis of the Worldwide Governance and Doing Business Indicators used for their aid-allocation decisions and asked what they can do to improve the ratings.

However, some developing countries' governments and scholars now increasingly tend to resist international governance ratings, according to the experience of some World Bank and UNDP staff members interviewed for this paper. They challenge the indicators for normative, legitimacy, and methodological reasons. Some developing countries express their distrust to indicators that they resent as pro-liberalisation and pro-market and therefore as a continuation of the Washington Consensus. They accuse governance indicators to reflect a Western or Anglo-Saxon type of governance ideal that is not helpful in understanding the governance reality in their countries (see also Girvan 2002).

Krishna Guha reports in the Financial Times (2007, a, b) that nine of the World Bank's executive directors, including those from China, Russia, Mexico, and Argentina, complained about the most recent WGI ratings: "Some governments argue that the bank is in no position to preach on governance, following its own crisis over former

president Paul Wolfowitz's handling of a secondment package for his girlfriend. Others, including China, question whether the World Bank should be involved at all in rating countries on metrics such as accountability.' (Guha, 2007b) The motivations for countries to criticise the ratings differ, according to some interviewees: some countries simply do not agree with the methodology and normative content of the indicators or believe that they fail to grasp the reality in developing countries while other countries use the ratings as a pretext to attack the whole governance agenda of the World Bank or to avoid discussions about their own governance problems.

Some World Bank staff were also worried that their governance indicators could induce countries to undertake the wrong reforms (see previous sections), in the absence of clear evidence that such governance reforms would actually trigger economic growth.

5.2.3 Enhance country dialogue

Some World Bank staff reported that the WGIs and/or the CPIA are harmful for their relations with local governments. To illustrate, World Bank staff lost credibility in a country because government officials were convinced that the WGI ratings reflected a lack of basic regional knowledge. Another illustration is the non acceptance of low CPIA ratings by developing country governments. This explains partly why countries' ratings are too high, according to some experts interviewed: a World Bank country director has both an incentive to have good relations with the government as well as to obtain resources

5. The politics of governance ratings

for the country. Both depend on a good CPIA rating on which the country director has an influence.

Other staff members interviewed found the CPIA and the WGI useful as an entry point for country dialogue. While one staff member interviewed explained that the WGIs are particularly useful in those countries where the government is not very willing to reform, others explained that they have seen more use as an entry point for country-dialogue in those countries which score well on the indicators.

Several World Bank economists who work in country offices said that they were unable to fulfil the requests of developing countries' governments to explain to them their ratings on the WGIs. Many of the staff of international organizations interviewed who are working directly in developing countries said they do not use broad governance indicators like the WGIs on the ground. Some staff felt that it is difficult to start a policy-oriented debate in their country on the basis of broad composite indicators and some resented the stigmatization of countries through international governance ratings. Both more specific cross-country comparable indicators as well as into-depth assessments in a country—for instance the World Bank Institute's Governance Diagnostic Surveys—were cited by international organizations' country office staff as useful for country dialogue. To illustrate, the World Bank Governance diagnostic surveys in Indonesia in 2001 helped to identify those governance problems that citizens in the country thought to be the most urgent to tackle. This gave an incentive for the government to reform these

areas in order to increase their popularity, according to a World Bank staff member. The steering committee consisted mainly of local stakeholders who were also in charge of communicating the results to the public. Staff interviewed also emphasized the large number of meaningful indicators available and suggested consulting country ratings on the sources directly: “If you want an indicator of how foreign experts assess the level of corruption, pick International Country Risk Guide or Economist Intelligence Unit. If you want an indicator of how local firms or households assess corruption, pick the World Economic Forum or another survey.” A staff member interviewed also highlighted the website of the WGIs as an excellent tool for gaining an overview of and accessing available sources of governance data.

5.2.4 Higher transparency of aid-allocation decisions

Donors are increasingly aware of the lack of good governance in their own aid allocation processes and, therefore, consider governance indicators a tool to enhance the objectivity and transparency of those decisions. Paradoxically, the indicators the World Bank uses and the World Bank indicators used by outside donors such as the MCC lack just that transparency (Arndt and Oman, 2006). To illustrate, the CPIA rating has only recently been made public. Historical data, ratings for the International Bank for Reconstruction and Development countries (i.e., middle-income

countries) and sub-ratings and detailed justifications produced by the staff undertaking the assessments are still not public.

5.2.5 Quantitative governance analysis

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

Yet researchers' use of governance indicators is widely plagued with pitfalls, as the nature and limitations of composite governance indicators unfortunately weaken the rigour and the credibility of many studies (Arndt and Oman, 2006).

According to a World Bank staff member interviewed, products such as the WGIs only allow for broad macro studies, with due attention paid to their conceptual and technical limitations. They cannot be used to identify sequencing issues in reform (e.g., do public expenditure reforms precede civil service reforms, etc.?) and apparent relationships (e.g., the relationship between improvements in public expenditure management and improved service delivery, etc.). Some

World Bank staff members interviewed expect that new generations of governance indicators currently under development will be more helpful for upstream research purposes.

5.2.6 Institutional reasons

A number of institutional reasons reinforced the rise of governance indicators in international organizations. First of all, governance indicators give organizations and divisions within organizations a profile. It is a product that is very visible within and outside the organization. Secondly, if a leading actor adopts a new policy, other actors are likely to follow independent of its efficiency (DiMaggio and Powell, 1991; Tolbert and Zucker, 1983). Some divisions within the World Bank produced and used governance indicators, and other divisions as well as other development agencies and international organizations like the OECD and the Asian Development Bank followed. They followed, because governance indicators are “trendy” and/or because they aimed at improving upon existing measurements or complementing them with different measurements. Thirdly, visible work on the increasingly important issue of measuring governance can help units within an international organization to obtain a more important role. Dunleavy’s (1991) bureau-shaping model predicts that bureaucrats would prefer to work in small, elite agencies close to political power centres and doing interesting work, rather than to run large-budget agencies with many staff but also many risks and problems. To illustrate, the popularity of the WGIs contributed to a gain in strategic importance of the World

5. The politics of governance ratings

Bank Institute and the Doing Business Indicators made the International Finance Corporation, 1998, (IFC) of the World Bank more prominent, according to some staff members interviewed. It is precisely the high level of usage and media coverage of their indicators that benefited the World Bank Institute and the IFC (see for instance World Bank Institute, 2008). Of course the WGIs are only a small part of the World Bank Institute's work (and the Doing Business Indicators are only a small fraction of the IFC's work), and changes in its status can therefore not be solely attributed to the WGIs. Furthermore, some of the criticism of the WGIs comes from within the World Bank Institute.

The increase in the number of producers of indicators has had a positive effect on the quality and diversity of governance indicators available. Many meaningful datasets on the perceptions of households, firms, civil servants, and experts on a plethora of governance topics are available. To illustrate, in response to a previous lack of data, several new initiatives (e.g., DIAL's household surveys, see Razafindrakoto and Roubaud, 2006) produce indicators that reflect the experiences of the poor and other initiatives try to produce more transparent indicators (e.g., Global Integrity).

However, it is difficult for new initiatives like Global Integrity and DIAL to gain attention, because the most-widely used indicators are well-established and dominate the market. The success of the most-widely used indicators has created interests which make it more difficult to move on to new developments (see DiMaggio and Powell,

1991; Stinchcombe, 1965; Zucker, 1977). In other words, despite the efforts of new producers of indicators to promote more meaningful and transparent metrics, the present institutional setup, which was created by dominant players within the World Bank, renders difficult the widespread use of indicators alternative to the dominant ones. Over the last five years, the use of these dominant indicators outside of the World Bank has been institutionalised and the switching costs for users wishing to use alternative indicators can be sometimes prohibitive (see DiMaggio and Powell, 1991). The rationale behind the use of the dominant indicators will be exposed in the next section.

The downside of the “measuring governance trend” is that organizations spend large amounts of money on assessing governance, while the funds might, in some cases, be better used elsewhere. Jak Jabes (2002), Director of the Governance and Regional Cooperation Division at the Asian Development Bank, cautioned: “Governance indicators are ultimately useful if they can be utilized to mobilize the leaders of a country to bring about change. If, as I unfortunately think is happening, indicators become an end in themselves, then they end up becoming part of dust gathering reports and detract attention from a less complicated yet more practical and productive approach” (p. 11).

5.3 An attempt to explain the dominance of the Worldwide Governance Indicators over more meaningful indicators

Of all the indicators produced by the World Bank, the WGIs are the most popular among international investors, the media, bi- and multilateral donors, and development analysts and academics. Within the World Bank, their limitations are well acknowledged and staff are advised to use a range of indicators, taking into account their respective limitations: “During the recent consultations, the majority of stakeholders endorsed an approach of using a mix of indicators—aggregate governance indicators, country monitoring and evaluation systems, specific disaggregated indicators, and outcome indicators—judiciously and with care, keeping in mind their strengths and limitations for particular purposes” (World Bank, 2007, p.9).

Why are the WGIs so popular externally, despite their shortcomings, the scepticism and reluctance to use them within the World Bank itself, and the availability of many more meaningful and transparent cross-country comparable indicators (see IADB, 2007)?

The following four major reasons lie behind the market domination of the WGIs: 1) users do not understand the limitations of the indicators; 2) complex composite indicators meet an important demand for summary measurements of the quality of governance with a broad country coverage; 3) users follow other users; and 4) users perceive that there is no alternative. The main focus of this section is on the World Governance Indicators and not on the Doing Business

Indicators, because the latter only concern one aspect of governance and lend themselves therefore to fewer uses than the WGI.

5.3.1 Incomplete understanding of indicators' limitations

Users do not fully grasp the limitations of the indicators they use and have therefore wrong perceptions of the reliability and content of indicators. In particular, they are often not aware of problems with the comparability over time and across countries, and of the normative content of indicators. (To illustrate, a recent paper from a renowned university refers to the WGI and Freedom House indicators as “objective”.)⁵⁹ The reasons for this misunderstanding are multiple and intertwined.

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

⁵⁹ Johnson and Zajonc (2006) refer to the indicators used for the MCC (among them several perception-based indicators like the WGI and Freedom House's Political Rights and Civil Liberties indicators) as “a set of objective and transparent governance indicators.”

5. The politics of governance ratings

impressed by the equations and tests in the background papers that they do not fully understand.

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

Thirdly, non-statistical experts often tend to believe that numbers are facts. As soon as a governance assessment is quantitative, the assessment is believed to be objective. Numbers reassure and create the impression of indisputable facts. This is reinforced by the fact that the dominant indicators are produced by a leading organization, which is perceived as a high-status producer of scientific works. The quality of its contribution is hence not only often taken for granted, but also difficult to question (or questioning is seen as at best an “unnatural” task).

Fourthly, misleading advertising contributes to a considerable misperception and therefore misuse of governance indicators. While Transparency International carefully advertises its CPI as “a snapshot of the views of business people and country analysts, with less of a focus on year-to-year trends,” the World Bank Institute advertises its WGI as “reliable measurements of governance.” Presentations and its website also create the misleading impression that the views of ordinary citizens are well represented, making the indicators

particularly attractive to donor agencies that seek to improve the living conditions of the poor. To illustrate, Kaufmann, Kraay, and Mastruzzi (2007a) heavily emphasize the recent inclusion of the source “Gallup World Poll,” a cross-country household survey available for a large number of countries. What they do not emphasize in the paper is that Gallup’s World Poll gets zero weight on two WGI, marginal weight on two other WGI, and provides no data for the remaining two.

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

5.3.2 Supply meets demand for summary measurements with broad country coverage

There is a high demand from development agencies, ministries, international organizations, international banks, and the media for summary measurements of the quality of governance with broad country coverage. Summary measurements are convenient for busy decision makers. They are quick to look at and therefore help to make

5. *The politics of governance ratings*

quick decisions, compared to looking at hundreds of different indicators or qualitative reports.

Pollitt (forthcoming) explains the prominence of the WGIs by the “power of numbers” and their prominent role within “late modernism,” citing the works of Alonso and Starr (1987), Porter (1995), Tsoukas (1997), Bevan and Hood (2006), Boyne *et al.* (2006), Bouckaert and Halligan (2008), Hood (2007), and Van de Walle and Roberts (2008). Van de Walle and Roberts (2008) suggest that ratings may serve to provide the illusion of control in an increasingly complex and turbulent world.

Since Kaufmann and colleagues include a country in their rating as long as at least one source provides data for it, the number of countries the WGIs provide data for is very large compared to other indicators. Some researchers interviewed for this study were aware of some of the limitations of the WGIs, but preferred to ignore them in order to maintain a larger number of countries in their dataset.

Using a summary measurement of most available sources frees decision makers and development researchers from getting lost in the complexity of governance and from making explicit normative choices that could render them vulnerable to criticism. To illustrate, Van de Walle (2008) shows that concepts used in evaluating the performance of public services appear to be neutral, but are generally value laden. Protecting public-service values such as universal access to services, including remote areas, can come at the sacrifice of

efficiency, and there is no neutral answer to what is more important. Similarly, most donors or academics agree that the quality of regulations should be good. The disagreement arises as soon as they try to specify what regulations are a sign of good quality: are environmental regulations deemed “good” because they protect the environment and the health of the population, or should they be considered to be “bad” because they put businesses under too much financial and administrative strain to comply with them? Most people also agree that governments should protect the freedom of the press, but they may disagree on the specific weight that different aspects of press freedom should carry in a “Press Freedom Index.” To illustrate, what weight should be given to an indicator assessing the number of journalists who were killed in a given country, over an indicator that assesses ownership concentration in the media sector?

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

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

5.3.3 Media

The media extensively use and publish ratings, because ratings sell. The American newspaper U.S. News & World Report, for instance, generates a lot of publicity over its annual ranking of colleges and universities and continues to publish the ratings, despite protests of American universities about the arbitrariness of these indicators. The controversy is similar to the controversy over the CPI and the WGI.

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

New initiatives like Global Integrity, which provides transparent, meaningful and actionable data on the existence and effectiveness of anti-corruption mechanisms, face not only an uphill battle against other indicators that have been around longer, but also against indicators that are available for larger number of countries. As a World Bank economist phrased it, “a better or more meaningful index can be crowded out by a less meaningful one that has broader country coverage.”

Political reasons may also explain why some indicators are more promoted and passed on to the media. To illustrate, in the field of corruption, the WGI “Control of Corruption,” which focuses on bribery and other illegal corrupt activities, is professionally promoted. It is interesting that the World Bank Institute’s work on legal corruption (Kaufmann, 2004, Kaufmann and Vicente, 2005) is barely promoted and therefore barely known. It takes into account legal ways (lobbying or party-financing) for special interest groups to influence laws and regulations in a country. These are more common in OECD countries than bribery to influence the government in favour of special interests. The Corporate Legal Corruption Index ranks the United States, a powerful and highly influential member of the World Bank, only 53rd in 2004 (the lower the rank, the better), while it scores among the top countries on the WGI “Control of Corruption” 2004.

5.3.4 Herd behaviour

Governance Indicators obtain authority (i.e., greater legitimacy) through the number and importance of users. Other users, therefore, often serve as a justification for someone’s own use. To illustrate, the authors of the IMF Country Report on Albania (IMF 2006) first describe the limitations of the WGI and then proceed to say, “However, the government (. . .) recognizes that the World Bank governance indicators are generally accepted internationally.” This creates a self-reinforcing mechanism that increases the taken-for-grantedness of these indicators, their mechanistic use, and in the end a bandwagon effect whereby more and more users adopt indicators

5. The politics of governance ratings

produced by leading organizations (e.g., Ellickson, 2001; Rao, Monin and Durand, 2003, 2005; Tolbert and Zucker, 1983).

5.3.5 Perceived absence of an alternative

Many users perceive that there is no alternative to the most widely used governance indicators. The reason is threefold: first of all, many alternative governance indicators are only available for a limited number of countries, whereas indicators such as the WGI draw their utility precisely from the fact that they use all available information to obtain a rating for almost all countries in the world. Secondly, users are likely to think that the more sources a composite indicator uses, the more reliable its rating will be. Thirdly, users may not even know about innovative approaches to quantify the quality of governance, because the media focuses its attention on established indicators like the WGIs, the Doing Business Indicators and Transparency International's Corruption Perceptions Index.

5.4 Conclusion and recommendations

The World Bank is the major producer of governance indicators, of both aggregated summary measurements (e.g., the WGIs), firm surveys (e.g., BEEPs) and more recent, “actionable” indicators (e.g., PEFA). The World Bank produces governance ratings to help identify areas of reform and measure the success of governance reforms, create incentives in developing countries to improve governance and put

governance on the agenda, to facilitate country dialogue, to enhance the transparency and objectivity of aid-allocation decisions, to enhance quantitative governance research and for institutional reasons.

The indicators which are most popular externally are hotly debated within the World Bank and their limitations acknowledged. Hence, within the World Bank, the WGIs for instance are mainly used to gain a first idea of how experts perceive a country's quality of governance and as an entry point for country dialogue, but not for more far-reaching decisions and judgments about countries' quality of governance.

However, outside the World Bank, the WGIs are also widely used for research, aid-allocation decisions, country risk analysis and media articles that praise or criticize a government—with too little attention paid to the indicators' limitations and despite the availability of more meaningful and robust indicators. The high demand for summary measurements of the quality of governance with a broad country coverage, lacking understanding of the limitations of the indicators, herd behaviour and a perceived lack of alternatives go far to explain the domination of the WGIs.

There are several alternatives to relying heavily on a few summary measurements of the quality of governance. In the short run, users can find advice and guidance about the measurements that will most likely suit their purposes in governance indicator guides and inventories (IADB, 2007; UNDP, 2007; Metagora, 2007; World Bank

5. The politics of governance ratings

Institute, 2007; Besancon, 2003). These are useful to gain awareness of and profit from the wealth of existing data and to understand the indicators' limitations and hence use them correctly. The Inter American Development Bank's (IADB) web tool, for example, provides information for 400 indicators about the methodology used to build them and the implications this has for the indicators' reliability, validity, and suitability for making comparisons across countries and over time. In the short run, producers can also improve the way they advertise their indicators and improve their transparency to better convey to users the meaning and limitations of their indicators.

In the medium-to-long run, the trend towards more specific, meaningful and actionable indicators will continue. New, more specific and transparent initiatives have emerged in recent years. Some provide comparable data for a large number of countries and others are more context-specific. What they have in common is that they do not claim to assess the quality of overall governance, but focus on specific and well-defined aspects of governance and, therefore, give guidance to developing countries on how to improve their rating. One prominent illustration is the PEFA initiative which measures critical dimensions of an "open and orderly public financial systems." Findings for Ghana in 2006, for instance, indicated that while comprehensive information is available on the intended use of public resources, in-year reporting on the utilisation of those resources is less comprehensive, hampering the monitoring of budget performance and reducing the efficient management of overall budget

operations. The PEFA secretariat started in September 2007 to monitor the actual use and follow-up of reports, but it is too early to judge its actual impact. Another illustration is the pipeline World Bank project “LAC Middle Income Governance and Public Management Partnership Facility” which plans to develop comparative indicators in the field of public management for middle-income Latin American countries. Until we have at our disposal a theory of governance able to guide the construction of meaningful indicators of the overall quality of governance, the development community should focus on such specific and well-defined indicators. As with all indicators, these new indicators have their limitations; however, since they are transparent, their limitations are also transparent, hence reducing the danger of misinterpretation and misuse.

Quantifying the quality of governance in a transparent and specific way and trying to make indicators “actionable” forces development analysts to face the limits of their knowledge. One major motivation for international organizations to develop “actionable” indicators is to induce governments to reform in order to improve their rating on these indicators. Developing indicators that suggest specific reforms requires a deep understanding of which governance reforms are conducive to development. According to some development experts interviewed, the lack of consensus on which governance reforms are necessary in developing countries and to what extent these reforms are context-specific, might precisely explain the current lack of many cross-country comparable actionable indicators.

5. The politics of governance ratings

Some experts expressed their fear that an overambitious good governance agenda, reflected in a broad set of good governance indicators, might render a developing country's government immobile and divert its attention from more urgent problems. Several development experts argued for more context-specific governance assessments that support governments in setting priorities for governance reform and monitor progress over time (e.g., World Bank Institute Governance Diagnostic Surveys, UNDP governance assessments).

Finally, some World Bank staff, interviewed, expressed the need for more debate and research, both more qualitative work and quantitative work based on more meaningful and robust indicators, to better understand developing countries' governance systems and the governance-growth nexus (e.g., Meisel, 2004). It is much better to have this debate now than to start (as happened with so many other development fashions) implementing governance reforms on a large scale that, despite good intentions, might hinder country development more than help it.

6 Rule of Law and economic growth: A panel data re-examination⁶⁰

⁶⁰ This chapter is a working paper (Arndt and de Crombrugghe, 2008).

Over the past decade, the focus of development policies shifted away from mainly macroeconomic concerns towards improving governance. The World Bank, for instance, has shifted its priorities from lending for economic reforms (which fell by 14 per cent annually between 2000 and 2004) to improving governance (which rose by 11 per cent). North's (1990, 2005) work was crucial to provide the theoretical underpinning for this fundamental change in development practice: He demonstrated the importance of a country's system of governance for the country's success in terms of its long-term economic growth. A series of quantitative studies followed, providing some empirical evidence of a positive impact of good governance on economic performance and on the effectiveness of aid⁶¹.

Nevertheless, the exact nature of the governance-growth relationship remains a subject of ongoing debate in the empirical literature. The standard problem in identifying the effect of governance on economic development is the endogeneity of governance, caused by simultaneity, measurement error and/or omitted variables. Mauro (1995) attempted to correct for reverse causality by constructing exogenous instrumental variables for his governance index, namely an index of ethnic fractionalization and a set of colonial heritage dummies. The exogeneity of these instruments

⁶¹ Classic references on the empirical relationship between governance and economic performance are, e.g., Acemoglu, Johnson and Robinson, 2001; Kaufmann and Kraay, 2002b; Knack and Keefer, 1995; Rodrik, Subramanian and Trebbi, 2002. See Aron, 2000, Knack, 2002, 2006, and Gray, 2007 for literature reviews. On the relationship between governance and aid effectiveness see, e.g., Burnside and Dollar, 1997, 2004.

6. *Rule of Law and economic growth: A panel data re-examination*

has been questioned (see Easterly and Levine, 1997; Zak and Knack, 2001). Acemoglu *et al.* (2001) use settler mortality as an instrument for institutions, which was picked up by a series of other authors (e.g. Kaufmann and Kraay, 2002b, Rodrik *et al.*, 2002). The finding by Acemoglu *et al.* of a strong positive impact of institutions on long-run economic growth was heavily debated: Sachs argued that institutions are almost entirely endogenous to growth (see Sachs and McArthur, 2001; Sachs and Warner, 1995) and Glaeser *et al.* (2004) demonstrated that settler mortality may simply capture the human capital of the European colonists, rather than early institutional strengths. Albouy (2006) found that Acemoglu *et al.*'s results are highly sensitive to the choice of information sources on settler mortality, with multiple mortality estimates being available for several cases. Rigobon and Rodrik (2004) avoided the delicate search for instrumental variables altogether and proposed to replace it by a method for achieving identification through heteroskedasticity. Still, with their new technique they did find, in line with the existing literature, that the Rule of Law has a strong positive effect on income levels.

Many of those studies investigating the effects of governance on economic performance focussed on per capita income *levels* rather than growth. Whereas the existence of some long-run causal link from governance to income seems to be widely accepted, its exact nature is still under debate. In particular, according to some authors, no clear evidence has been produced as yet that governance impacts positively on economic growth for the policy-relevant time horizon of one or

two decades. Rodrik (2008, p. 19) states that he is “not aware of any strong econometric evidence that relates standard governance criteria to growth (all the evidence is about income levels)”. Glaeser *et al.* (2004) argue that in their growth regressions human capital is a much more important determinant than are institutional factors.

Studies that do focus on the effects of governance on economic growth tend to use cross-sectional rather than panel data. One exception is Chong and Calderon (2000), who exploit the variation over time in growth and governance data to investigate the issue of causal feedback. Applying a version of “Granger causality” testing, they find strong evidence of causality in both directions.

Questioning the robustness of cross-country growth regressions is not unprecedented. Predating the governance-growth debate, Levine and Renelt (1992) demonstrated how sensitive growth regressions can be to small changes in the set of conditioning variables. Sala-i-Martin (1997) followed suit with two million regressions, and his somewhat more lenient criteria produced more positive conclusions. Kalaitzidakis, Mamuneas and Stengos (2000) examined the sensitivity of growth regressions to the introduction of some non-linear effects. We investigate here whether recently accrued data sets, panel data methods and new insights from the quality of governance debate help resolve the fragility of growth regressions.

Fragility can arise from many sides. Levine and Renelt (1992) and Sala-i-Martin (1997) concentrated attention on the choice of

6. Rule of Law and economic growth: A panel data re-examination

conditioning variables. In this paper we bring together different sources of data on the quality of governance and focus on the choice among several available proxy indicators for the Rule of Law as well as the way they are scaled. We concentrate on the Rule of Law for the following reasons: i) The Rule of Law is frequently used as a proxy for governance in econometric studies (e.g. Kaufmann and Kraay, 2002b; Rigobon and Rodrik, 2004), ii) the role of the Rule of Law in economic development has become “the motherhood and apple pie of development economics” (The Economist, 2008). Already Douglas North (1990) argued that impersonal exchange with third party enforcement is “the critical underpinning of successful modern economies involved in the complex contracting necessary for modern economic growth” (North, 1990, p.35).

The contribution of this paper is threefold. First, we rely on a panel data model with a time-recursive structure to identify the effect of governance on economic growth over the policy-relevant periods of one or two decades. In the past such work has been more difficult to conduct because of the scarcity of panel data on governance providing some historical depth. In spite of interesting attempts it is practically hopeless to establish the simultaneous causal relationships between growth and governance in a static cross-sectional setting, for want of generally convincing instrumental variables. The main difficulty, as Brock and Durlauf (2001) put it, is the open-endedness of growth theories. There exist so many mutually compatible theoretical arguments that it is very hard to justify exclusion restrictions. Hence, practically any candidate instrument is likely to be correlated with

potentially relevant omitted variables, thereby violating the exogeneity requirement. One example of this is Glaeser *et al.*'s (2004) argument about Acemoglu *et al.*'s (2001) settler mortality standing in as an instrument for human capital rather than for institutions as intended. The longitudinal (time) dimension in panel data provides another perspective, improving the prospect for formulating plausible identifying exclusion restrictions. The relevance of the governance-growth linkage is not with respect to short-term fluctuations but rather with respect to medium to long-term performance in economic growth. It seems to us reasonable to expect a time lag of a few years (say of the order of 5) for changes in governance, institutions and procedures to seep through economic activity and exert their impact on growth. Similarly, any real feedback from economic prosperity towards institutions and procedures of governance will take time to materialise and be implemented. Admittedly, the perceptions of experts and survey respondents about the quality of governance are likely to be influenced much more rapidly by economic conditions in their country. By the same token, any instantaneous effect of growth on governance indicators is more likely to be the result of these perceptions than evidence of a real feedback or substantive instantaneous changes in the institutions and procedures of governance.

Our second contribution is the compilation of a comprehensive collection of existing Rule of Law indicators and other variables for the largest possible set of countries. This will form the basis for tests of the fragility of results depending on the choice of the proxy indicator included in the growth regression. As indicated above, the

6. Rule of Law and economic growth: A panel data re-examination

sensitivity analyses of Levine and Renelt (1992) and Sala-i-Martin (1997) were based on varying the choice of conditioning variables rather than the choice of governance indicator.

Thirdly, we attempt solutions to a number of econometric problems likely to be present. (a) We deal with heterogeneity in the form of unobserved effects, allowing for correlation with regressors in the spirit of Mundlak (1978). We also consider certain forms of nonlinearity or heterogeneity in slope coefficients (e.g. allowing for interactions involving the governance indicators). (b) We treat the endogeneity due to measurement errors in initial income by standard instrumental variables methods. (Some omitted variables bias may also be captured on the way.) (c) Since governance indicators really reflect qualitative rankings, although they are mostly presented on pseudo-metric scales, we avoid cardinal interpretations as well as linearity assumptions about their relationship to growth. We transform the published governance indicators into categorical variables (dummies) indicating whether a country's quality of governance is rated markedly high (from the upper quartile up) or markedly low (from the lower quartile down) as compared to the median pack (those countries rated between the lower and upper quartiles).

The structure of the remainder of the chapter is as follows.

- Section 6.2 presents the standard growth model on which our specifications are based.

- Section 6.3 describes the database and the main variables used, concepts of Rule of Law, and the correlations found among them.
- Section 6.4 discusses specification issues and the sequence of choices to be made regarding not only the estimation method but also the source of the indicator data, the sample period used, the data periodicity, the specification of the indicator itself as a pseudo-metric or a qualitative, categorical variable, the set of regressors or control variables used, the treatment of income for its potential endogeneity.
- Section 6.5 presents empirical results in particular as they exhibit the fragility of (the significance of) the effects found as we vary the choices made. In Subsection 6.5.1 we estimate our baseline growth equation, excluding any Rule of Law indicators for the time being, yet varying the choice of method and of sample period. We present results from Pooled OLS, random effects and fixed effects regressions (Tables 6.7 and 6.A.01-6.A.04), as well as the analogous 2SLS regressions treating the initial income level as endogenous (Tables 6.8 and 6.A.05-6.A.08). In Subsection 6.5.2 we extend the regression equations with the four main competing Rule of Law indicators, one by one. We present summary tables on the direction and significance of the governance effects found. In Subsection 6.5.3 we discuss the fragility of the estimated coefficients of control variables to the choice of Rule of Law indicators. In Subsection 6.5.4 we investigate whether investments in physical capital

6. Rule of Law and economic growth: A panel data re-examination

might be themselves determined by and thereby masking the indirect effect of the Rule of Law.

- Section 6.6 concludes with a summary of findings, caution, practical meaning of results for reform advice, recommendations for future research.

6.1 Growth model

This section presents the standard growth model on which our specifications in the sequel will be based, and discusses how to deal with the heterogeneity in a panel of countries from all continents. Related issues of interest include the income-level/growth dilemma and how it relates to the neoclassical vs. endogenous growth debate; the so-called convergence debate; and time series aspects.

Neoclassical growth models (à la Solow-Swan) predict parallel steady state growth paths for the living standards (per capita incomes and outputs) of all the world's countries. The universal rate of growth in the steady state is exogenously determined (possibly by means of a time trend) by worldwide technological progress. The common trend does not imply a common standard of living, though. Output and income per capita in the steady state depend on the rate of saving or capital accumulation, possibly for distinct categories of capital (physical and human, private and public), as well as on population growth (which dilutes the capital intensity). To the extent countries maintain different saving habits, investment policies and/or fertility

rates, disparities in standards of living will also subsist, even if all growth rates converge to that of the worldwide technological trend.

It is possible to assume that the technology of social organisation, like the technology of production, follows some exogenously driven trend and, accordingly, to subsume the quality of institutions or governance into “total factor productivity”, the technological component of a neoclassical growth model. However, the plausibility of such an exogenously driven trend in the technology of social organisation is a matter for both theoretical and empirical debate. If countries can maintain different preferences toward savings, investment and fertility, it would seem *a fortiori* that they can perpetuate different tastes with respect to social organisation, and may therefore refuse to conform to some hypothetical worldwide trend in governance practices.

To facilitate the discussion we introduce some formal notation. Let $y_{i,t}$ represent the natural log of GDP (income or output) *per capita* for country i in time period t ; and $x'_{i,t}$ a vector of regressor variables thought to determine the steady-state output per person: saving or investment rates, labour force or population growth rate, technological time trend. As is conventional, the subscript i refers to the cross-section of countries ($i = 1, \dots, N$) and the subscript t to a fixed set of calendar periods (say, quinquenniums denoted by their closing date: $t = 1946, \dots, 2006$). Some empirical studies formalise the

6. *Rule of Law and economic growth: A panel data re-examination*

determination of living standards by regression equations of the following type:

$$(6.1) \quad y_{i,t} = x'_{i,t} \beta + u_{i,t}$$

where β is a vector of parameters and $u_{i,t}$ a random disturbance or error term. This can be interpreted as a description of the steady state path of per capita output as driven by factors like capital accumulation, work force growth and the technological trend. Deviations of a country from its steady state path are likely to be at least somewhat persistent over time, however, and therefore the disturbance term will be serially correlated. This has not been a concern in many studies estimating such equations using a purely cross-sectional data set. On the other hand, studies investigating per capita output as a single time series cast some doubt on the trend-reverting property it should have in a neoclassical world.

It is conceptually straightforward to extend the set of regressors with other variables possibly affecting either steady-state output or deviations from it: investment in human (as distinct from physical) capital (Barro 1991, Barro and Sala-i-Martin 2004, Mankiw, Romer and Weil 1992, Benhabib and Spiegel 1994), political stability (Alesina and Rodrik 1994), financial sector development (King and Levine 1992), trade openness (Frankel and Romer 1999), the quality of macroeconomic policy, governance or the quality of institutions (Knack and Keefer 1995, Rodrik *et al.* 2002), trust or social capital

(Knack and Keefer 1997), income inequality, and geography (Sachs 2003). For explicitness, let us have a vector of governance indicators $z'_{i,t}$, and replace the preceding equation by

$$(6.2) \quad y_{i,t} = x'_{i,t}\beta + z'_{i,t}\zeta + u_{i,t}.$$

The serial correlation in the disturbances is likely to be mitigated somewhat but not resolved by the addition of the institutional indicators. In fact, a major concern is that if the neoclassical model with its exogenous technological trend is incorrect, then deviations from the steady-state outcome might not even be stationary, rendering estimation of the relationship in the time dimension meaningless. (Cross-sectional regressions are not affected by this.)

Growth models are different in that they build in deviations from the steady-state path and are actually derived as a process of linear approximation around it. Let p denote the length of a standard period of time (say 5 years, in order to smooth out short-run business cycles as well as gaps in governance and other indicators). It will be convenient to denote annualised growth over a period $[t, t+p]$ subsequent to initial time t by

$$(6.3) \quad g_{i,t+p} = \frac{1}{p}(y_{i,t} - y_{i,t-p})$$

6. Rule of Law and economic growth: A panel data re-examination

A typical growth equation, that is, a regression relationship explaining growth by the initial output level, labour force and capital accumulation, a technological trend, and one or more governance indicators, might read as follows:

$$(6.4) \quad g_{i,t+p} = \alpha_i + y_{i,t}\beta + x'_{i,t}\gamma + z'_{i,t}\zeta + \varepsilon_{i,t}$$

The α_i represent country-specific trends (intercept heterogeneity); the $x'_{i,t}$ vector will in general include fixed time effects (period dummy variables) as well.⁶² The remaining terms $\varepsilon_{i,t}$ are “idiosyncratic” disturbances. In order to acknowledge the possibility of systematic heterogeneity in the slope coefficients, we may insert in addition interactions between (by way of example) one indicator, $z_{1,i,t}$, and the $x'_{i,t}$ vector:

$$(6.5) \quad g_{i,t+p} = \alpha_i + y_{i,t}\beta + x'_{i,t}\gamma + z'_{i,t}\zeta + z_{1,i,t}x'_{i,t}\xi_1 + \varepsilon_{i,t}$$

Much more general treatments of such parameter heterogeneity are possible.⁶³

⁶² In a one-time cross-section of countries, the individual country effects are absorbed in the disturbance term and there can be no time effects.

⁶³ Lee, Pesaran and Smith (1997) demonstrated the importance of allowing for slope heterogeneity, e.g.:

$$g_{i,t+p} = \alpha_i + y_{i,t}\beta_i + x'_{i,t}\gamma_i + z'_{i,t}\zeta_i + z_{1,i,t}x'_{i,t}\xi_i + \varepsilon_{i,t}$$

The growth equations above include unobserved heterogeneity in the form of time-constant country effects denoted α_i . Following the argument of Mundlak (1978), unobserved individual effects are likely in general to be correlated with observed characteristics in $x_{i,t}$ and $z_{i,t}$. The standard way to deal with this problem in linear models is a data transformation (either “within” or “first differencing”) designed to eliminate the unobserved effects from the model. The cost of such elimination methods, however, is that the effects of interest must be inferred from the within-country time series variation only; the cross-sectional variation is entirely differenced away, and the effects of variables that vary slowly over time are estimated imprecisely. Cross-sectional effects that are strictly invariant over time are entirely eliminated by differencing (although they can be reconstructed in another stage). This is in contrast with random effects methods that are based on the independence of α_i from $x_{i,t}$ and $z_{i,t}$. In order to dispose with this classical but unattractive independence assumption we allow for unobserved effects correlated with observed individual characteristics, in the following way. Let us denote time-averages, initial levels or similar time-constant values of the observed characteristics in $x_{i,t}$ and $z_{i,t}$ by so-called “Mundlak terms” \bar{x}_i, \bar{z}_i and interpret them as a set of persistent individual features.⁶⁴ We can

⁶⁴ Mundlak (1978) actually argued that in general all regressors should appear in the auxiliary regression and therefore ultimately the random effects model would be no different from the fixed effects model. We vary somewhat from Mundlak in that when it makes sense the time-constant values \bar{x}_i, \bar{z}_i may represent initial values x_{i1}, z_{i1} rather than averages taken over time.

6. Rule of Law and economic growth: A panel data re-examination

decompose the unobserved individual effects into a systematic part linearly dependent on \bar{x}_i , \bar{z}_i and an unexplained remainder term; in self-explanatory notation:

$$(6.6) \quad \alpha_i = \pi_0 + \bar{x}_i' \pi_1 + \bar{z}_i' \pi_2 + \eta_i$$

Substituting this expression for α_i , model (4) is rewritten as

$$(6.7) \quad g_{i,t+p} = \pi_0 + y_{i,t} \beta + x_{i,t}' \gamma + z_{i,t}' \zeta + \bar{x}_i' \pi_1 + \bar{z}_i' \pi_2 + \eta_i + \varepsilon_{i,t}$$

This last formulation of the model exhibits the composite disturbance with two “error components”, $\eta_i + \varepsilon_{i,t}$, implying a non diagonal variance-covariance structure. Efficient estimation in such a case requires a feasible Generalised Least Squares (GLS) method. Mundlak (1978) showed that GLS estimation of (6.7) with all terms included boils down to fixed effects estimation of (6.4); that is, if all regressors appear both in time-varying and in time-averaged form, the random effects model is no different from the fixed effects model. We will compare fixed effects estimation of (4) with several “shorter” regressions, where terms that vary slowly over time are included either in time-varying form or in time-averaged form but not both. We report two set of estimates, the first set treating $y_{i,t}$ as exogenous and the second instrumenting for $y_{i,t}$ using its lagged value. The first set comprises the following four estimates, in increasing order of generality:

1. Pooled OLS for equation (6.4), which we call “Pooled OLS”
2. Random effects GLS for equation (6.7) with every regressor appearing either in time-varying form or in time-averaged form but not both, which we dub “random effects with Time Averages” (RE1)
3. Random effects GLS for equation (6.4) (RE)
4. Fixed effects (FE) for equation (6.4), which amounts to random effects GLS for equation (6.7) with all terms included.

The second set consists of two-stage least squares (2SLS) analogues, treating the lagged income level as endogenous:

1. Pooled 2SLS for equation (6.4), which we call “Pooled IV”
2. Random effects IV-GLS for equation (6.7) with every regressor appearing either in time-varying form or in time-averaged form but not both, which we dub “random effects with Time Averages and IV” (RE1-IV)
3. Random effects IV-GLS for equation (6.4) (RE-IV)
4. Fixed effects IV (FE-IV) for equation (6.4), with IV for income⁶⁵.

In this second set of estimates, the list of instruments used is the list of regressors in the equation, minus $y_{i,t}$ and plus $y_{i,t-p}$, so in

⁶⁵ Fundamentally, Equation (6.4) is a dynamic panel data model and with a short time dimension fixed-effects methods (including FE-IV as used here) are subject to the so-called Nickell bias. Interesting alternatives for FE-IV estimation are bias-corrected formulae or methods based on first-differencing rather than the within-group transformation, in particular GMM and system-GMM. However, the scope of this study is exploratory and at this stage we prefer to emphasise specification choices rather than the technicity of estimation.

6. *Rule of Law and economic growth: A panel data re-examination*

effect $y_{i,t}$ is instrumented by its lag $y_{i,t-p}$. The pooled estimation ignores any form of heterogeneity and is most restrictive. The fixed effects estimation including the time-varying regressors only is the least restrictive, and can be used as a basis for comparison and for specification testing. The random effects estimates can be seen as a compromise between pooled and fixed effects estimation.

6.2 Data

We assembled a comprehensive data set, including many of the publicly available governance indicators as well as macroeconomic variables. The main sources are (i) the Quality of Government database of Teorell *et al.* (2008), (ii) the International Country Risk Guide from Political Risk Services (commercially available online), (iii) the Penn World Tables 6.2 of Heston *et al.* (2006), (iv) the World Bank's (2007) World Development Indicators, and (v) the geography data set of the Center for International Development at the Harvard Kennedy School. This section describes i) our data periodicity and scaling, ii) the choice of Rule of Law indicators and iii) the main control variables used.

6.2.1 Data periodicity and scaling

We draw on an unbalanced panel data set with yearly data from 1946 to 2006 for a total of more than 200 countries. We transform the periodicity of our dataset from yearly to five-yearly, covering

“quinquenniums” or five-year periods from 1947-1951 to 2002-2006. Within each quinquennium all stock and flow data are calculated as averages of the available yearly data (which conveniently reduces the proportion of missing observations). The panel is unbalanced since most variables are missing for some countries in some periods (especially in the early decennia). Accordingly, reported regressions draw on different numbers of periods and observations depending on the variables used, as indicated in the tables. Since these different sample coverages are not neutral for the regression outcomes, results will also be reported under a “common sample” restriction, using only countries and years for which all four governance indicators are available.

6.2.2 Rule of Law

The Rule of Law has become a “big idea in economics” (The Economist, 2008), heavily inspired by Douglass North (1990) who emphasized the importance of solid formal and informal institutions that reduce transaction and enforcement costs and hence foster economic growth. While there is much support in the academic community for the claim that Rule of Law is crucial to economic growth, the precise content of the “Rule of Law” is far from clear. Dani Rodrik wonders in his blog: “Am I the only economist guilty of using the term without having a good fix on what it really means?”⁶⁶ And Thomas Carothers (2003) from the Carnegie Endowment for

⁶⁶ http://rodrik.typepad.com/dani_rodriks_weblog/2007/11/what-does-the-r.html

6. *Rule of Law and economic growth: A panel data re-examination*

International Peace laments the lack of knowledge at many levels of conception, operation and evaluation surrounding the expanding “rule-of-law-promotion” in developing countries, an issue we will come back to in the conclusion. According to Trebilcock and Daniels (2008), the meaning of the Rule of Law for economists varies from very “thin” descriptions, which focus on property rights and the efficient administration of justice in order to provide for stability, to “thick” conceptualizations that treat the Rule of Law as the core of a just society, linking it to freedom and democracy.

Despite or perhaps because of these doubts about what the Rule of Law precisely is and more to the point what kind of Rule of Law would be necessary at different levels of development to stimulate economic growth, many indicators are available to assess whether a country has a well-functioning Rule of Law system.

We selected four Rule of Law indicators with sufficient country and time availability to allow for panel data analysis methods such as fixed and random effects to be used. The four indicators are i) the “Property Rights” indicator, which is a subcomponent of the Heritage Foundation’s Index of Economic Freedom (Heritage Foundation, 1995-2008); ii) the “Law and Order” indicator, which is a subcomponent of the International Country Risk Guide (ICRG), produced by the Political Risk Service (2007); iii) the composite indicator “Rule of Law”, which is part of the World Bank Institute’s Worldwide Governance Indicators (WGIs, Kaufmann *et al.*, 2007) and iv) the “Legal Structure and Security of Property Rights” indicators,

which is a subcomponent of the Fraser Institute's Economic Freedom rating (Gwartney and Lawson, 2006).

All four indicators focus either mainly or exclusively on the Rule of Law for the private sector and are therefore closer to the "thin" description. They are widely reported in the media and used in quantitative studies⁶⁷. Their actual content differs though: While the Heritage Foundation's "Property Rights" index is more narrowly defined with criteria related to the protection of private property and enforcement of contracts, the Worldwide Governance Indicator "Rule of Law" contains in addition many variables related to the incidence of crime. Table 6.1 lists the indicators, their definitions and their sources. The four indicators are very much interdependent: The WGI "Rule of Law" is based, among other sources, on both the Heritage Foundation "Property Rights" indicator and the ICRG "Law and Order" indicator. Both the Heritage Foundation and the World Bank Institute use data from the Economist Intelligence Unit for their indicators. Both "Rule of Law" and "Legal Structure and Security of Property Rights" are partly based on data from the World Economic Forum's Global Competitiveness Report; and the "Legal Structure and

⁶⁷ See for instance

<http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0,,contentMDK:21407860%7EpagePK:64168445%7EpiPK:64168309%7EtheSitePK:1740530,00.html>

for press coverage of the WGIs,

<http://www.freetheworld.com/2007/5EFW2007app2.pdf> for all the studies using the Fraser Institute indicators, including the "Legal Structure and Property Rights" indicator.

6. Rule of Law and economic growth: A panel data re-examination

Security of Property Rights” indicator draws both on the WGI “Rule of Law” and the ICRG indicator “Law and Order”⁶⁸.

⁶⁸ There is a wide body of literature that critically analyses the governance indicators and to which we return later.

Table 6.1 Rule of Law indicators

Name and Source	Scale and definition	Type of indicator	Availability
<p>“Property Rights”</p> <p>Source: Heritage Foundation, QoG Database*, (1995 to 2007)</p>	<p>Scale from 0 to 100 originally (the higher the better), rescaled from 0 to 10 in our calculations.</p> <p>This factor scores the degree to which a country's laws protect private property rights and the degree to which its government enforces those laws. It also accounts for the possibility that private property will be expropriated. In addition, it analyzes the independence of the judiciary, the existence of corruption within the judiciary, and the ability of individuals and businesses to enforce contracts. The less certain the legal protection of property is and the greater the chances of government expropriation of property are, the higher a country's score is. Originally, each country was classified in one of five defined categories. Recently, the number of categories was extended to 11. The scale ranges from 0 (worst category) to 100 (maximum degree of protection of property rights) We rescaled the scores from 0 to 10.</p>	<p>In-house Expert assessment; draws on other expert assessments provided by the Economist Intelligence Unit, the US State Department and the US Department of Commerce.</p> <p>Small changes in data sources over time.</p> <p>“Property Rights” is a subcomponent of the “Index of Economic Freedom”</p>	<p>Years: 1994-2006 Countries: 163**</p> <p>Publicly available: http://www.heritage.org/index; QoG database*,</p>
<p>“Law and Order”</p> <p>Source: Political Risk Services, International Country Risk Guide (2007)</p>	<p>Scale from 1 to 6 (the higher the better)</p> <p>“Law and Order“ are assessed separately, with each sub-component comprising zero to three points. The Law sub-component is an assessment of the strength and impartiality of the legal system, while the Order sub-component is an assessment of popular observance of the law. Thus, a country can enjoy a</p>	<p>In-house expert assessment</p> <p>“Law and Order” is a subcomponent of the International Country Risk Guide (ICRG) rating</p>	<p>Years: 1984-2003 Countries: 145**</p> <p>Not publicly available Information on http://www.prs-group.com/ICRG.aspx</p>

6. Rule of Law and economic growth: A panel data re-examination

	high rating – 3 – in terms of its judicial system, but a low rating – 1 – if it suffers from a very high crime rate / if the law is routinely ignored without effective sanction (for example, widespread illegal strikes).		
WGI “Rule of Law” Source: World Bank Institute (2007), Worldwide Governance Indicators, taken from QoG Database*)	<p>Scale from -2.5 to + 2.5 (the higher the better)</p> <p>“Rule of Law” includes many indicators provided by 24 sources which measure the extent to which agents have confidence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts. Together, these indicators measure the 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 (Kaufmann <i>et al.</i>, 2006). Underlying indicators include the ICRG “Law and order” indicator and the Heritage Foundation “Property Rights” indicator.</p> <p>The governance estimates are postulated to be normally distributed and are normalised to have a mean of zero and a standard deviation of one in each year of measurement. This implies that virtually all scores lie between –2.5 and 2.5, with higher scores corresponding to better outcomes.</p>	Composite indicator, based on indicators from 24 expert assessments and surveys (number and composition of underlying indicators vary across countries and over time; a country is included if at least one source provides data)	<p>Years: 1996-2006 Countries: 202**</p> <p>Publicly available (data is not available for all subcomponents of “Rule of Law”) www.govindicators.org; QoG database*</p>
Fraser Institute “Legal Structure and Security of Property	<p>Scale from 0-10 (the higher the better)</p> <p>The subcomponents are (A) judicial independence, (B) impartial courts,</p>	In-house expert assessment, draws on the Global Competitiveness Report from the World Economic	<p>Years: 1970-2006 Countries: 129**</p> <p>Publicly</p>

<p>Rights” Source: Gwartney and Lawson (2006)</p>	<p>(C) protection of property rights, (D) military interference in Rule of Law and the political process, (E) integrity of the legal system, (F) legal enforcement of contracts, (G) regulatory restrictions on the sale of real property.</p> <p>The index ranges from 0-10 where 0 corresponds to ‘no judicial independence’, ‘no trusted legal framework exists’, ‘no protection of intellectual property’, ‘military interference in Rule of Law’, and ‘no integrity of the legal system’ and 10 corresponds to ‘high judicial independence’, ‘trusted legal framework exists’, ‘protection of intellectual property’, ‘no military interference in Rule of Law’, and ‘integrity of the legal system’.</p>	<p>Forum, the Worldwide Governance Indicators, the ICRG and the Doing Business Indicators.</p> <p>Sources differ across countries and over time</p> <p>“Legal Structure and Security of Property Rights” is a subcomponent of the Fraser Institute’s “Economic Freedom Index”</p>	<p>available (subcomponents of « Legal Structure and Security of Property Rights » only since 1995) http://www.freetheworld.com/ ; QoG Database*</p>
---	--	---	---

* Teorell, Jan, Sören Holmberg and Bo Rothstein. 2008. The Quality of Government Dataset, version 15May08. University of Gothenburg: The Quality of Government Institute, <http://www.qog.pol.gu.se>

** Number of countries for which there is data for at least one year. The number of countries for which there is data in any given year, in particular for historical data, can of course be lower.

6. Rule of Law and economic growth: A panel data re-examination

Table 6.2 shows correlations between the four indicators, Panel A in levels and Panel B in first differences (changes over time). Although the correlations are substantial, they also suggest that the indicators contain some independent variation. The highest correlation in the levels is between the WGI “Rule of Law” and the Fraser Institute’s “Legal Structure and Security of Property Rights” indicator, the lowest between the Heritage Foundation’s “Property Rights” indicator and the ICRG “Law and Order”. The correlation between changes over time in the indicators is, as one would expect, lower than the correlation between their levels. The highest correlation is between changes in the ICRG’s “Law and Order” indicator and the Fraser Institute’s “Legal Structure and Security of Property Rights” indicator and the lowest correlation is between changes in the latter and the Heritage Foundation’s “Property Rights” indicator.

Table 6.3 gives an overview of the strengths and weaknesses of the four indicators. Imperfect measurement is more than likely to be an unavoidable feature of all such constructed indicators. One specific problem occurring repeatedly with Heritage Foundation and ICRG is that a change in the rating from one year to the next does not always reflect an actual change in the quality of governance over time but is due to a non-retrospective correction or revision by the data producing agency. The occurrence of such corrections seems to suggest that unlike classical measurement errors which are assumed independent, observation errors in the Rule of Law indicators may be persistent and serially correlated. Another specific problem characteristic of the WGI composite indicator is its year-by-year normalization and relative

scaling. In effect, the scaling is not absolutely comparable over time and the outcome should be interpreted as a repeated cross-section rather than a longitudinal data set; it contains little or no time series information. Strictly speaking, it provides ordinal rather than cardinal information, and therefore has a pseudo-metric character. We will in the following sections try to alleviate the consequences of measurement errors, non-metricity and implied asymmetries or nonlinearities in the published indicators by a categorical transformation. For a critical review of these and other problems with governance indicators, see for instance Arndt and Oman (2006); Charles *et al.* (2008); Davis (2004); Glaeser *et al.* (2004), Knack (2007); Kurtz and Schrank (2007); Thomas (2007); van de Walle (2005). Davis (2004), in particular, undertakes an into-depth analysis of Rule of Law indicators (including the ICRG “Law and Order” indicator and the Heritage Foundation “Property Rights” indicator) and concludes that the indicators commonly used for quantitative analysis do not capture information capable of shedding light upon the potential impact of purely legal reforms, an issue we come back to in the conclusion. See Kaufmann, Kraay and Mastruzzi (2007b) for a reply to some of the comments on the WGIs.

Glaeser *et al.* (2004) criticise that most indicators used in studies to proxy for institutions measure “outcomes, not some permanent characteristics that North refers to” (p. 8). We indeed interpret our indicators as outcome measurements, and do not claim that they reflect the permanent quality of institutions. Instead, we explicitly interpret the indicators as proxies for one aspect of the quality of

6. Rule of Law and economic growth: A panel data re-examination

governance being subject to evolution and change. Indeed, common definitions of the quality of governance are perfectly compatible with changes over a time-horizon of five years⁶⁹.

Table 6.2 Correlations between different Rule of Law indicators			
Panel A: Correlations between levels			
	“Property Rights” Heritage Foundation	“Law and Order” ICRG	“Rule of Law” WGI
“Law and Order” ICRG	0.695		
“Rule of Law” WGI	0.878	0.831	
“Legal Structure and Security of Property Rights” Fraser Institute	0.806	0.860	0.929
Note: Common sample of 325 observations			
Panel B: Correlations between changes over time			
	“Property Rights” Heritage Foundation	“Law and Order” ICRG	“Rule of Law” WGI
“Law and Order” ICRG	0.090		
“Rule of Law” WGI	0.335	0.366	
“Legal Structure and Security of Property Rights” Fraser Institute	0.083	0.377	0.230
Note: Common sample of 212 observed first differences			

⁶⁹ For an overview of definitions of governance, see <http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/MENAEXT/EXTMNAREGTOPGOVERNANCE/0,,contentMDK:20513159~pagePK:34004173~piPK:34003707~theSitePK:497024,00.html>.

Table 6.3 Strengths and Weaknesses of Rule of Law indicators

	+	-
Heritage Foundation “Property Rights” Index	<ul style="list-style-type: none"> • Clear definition of categories • Explanation in handbook for every country’s rating and its changes • Sources similar across countries, few changes over time in sources • High country coverage (more than 150 countries) 	<ul style="list-style-type: none"> • Most changes of ratings over time are revisions due to new information or source corrections, not actual changes in the protection of property rights • Biased towards Anglo-Saxon system
ICRG “Law and Order”	<ul style="list-style-type: none"> • Available since 1984 (on average for 129 countries) 	<ul style="list-style-type: none"> • Data not publicly available • Detailed explanation of criteria used to rank countries not available, no explanation for changes • Some changes of ratings over time are revisions due to new information or source corrections, not actual changes in law and order enforcement
WGI “Rule of Law”	<ul style="list-style-type: none"> • Very high country coverage (cross-country: over 200, over time on average 185) 	<ul style="list-style-type: none"> • Conceptually weak • Serious lack of comparability over time and across countries
Fraser Institute “Legal Structure and Security of Property Rights”	<ul style="list-style-type: none"> • Available since 1970 (on average for 104 countries) • All disaggregated data available 	<ul style="list-style-type: none"> • Conceptually weak • Biased towards Anglo-Saxon system • Lack of comparability across countries and over time: Significant changes in sources and content over time

6.2.3 Dependent variable and other explanatory variables

The dependent variable to be modelled is the annual average growth rate in GDP per capita from the Penn World Tables 6.2 over a period of five years. The Penn World Tables are known for providing an up-to-date and comprehensive set of countries representative of the whole world.

We decided to control for regime type, using the classification originally developed by Przeworski, Alvarez, Cheibub, and Fernando (2000) and extended by Cheibub and Gandhi (2004). The advantage of the classification is threefold: i) it is transparent and replicable, ii) data is available for a maximum of 198 countries since 1946 and iii) there is no conceptual overlap with our Rule of Law measures. The Cheibub and Gandhi regime type variable is highly correlated with other democracy variables.

The variable is coded 0 if the country's regime is classified as a democratic, 1 if it is classified as a dictatorship. A regime is considered a dictatorship if the chief executive is not elected, the legislature is not elected, there is no more than one party, or there has been no alternation in power (Przeworski *et al.* 2000). Transition years are coded as the regime that emerges in that year.

The regime classification can be clearly distinguished from the Rule of Law indicators. Property rights protection and a low crime rate are possible in dictatorships and one-party states. A prominent

example is Singapore: Singapore is ranked 1st together with other countries on the Heritage Foundation's "Property Rights" indicator, scores among the top ten on the Fraser Institute's "Legal Structure and Security of Property Rights" indicator, has the second highest score next to countries such as Switzerland and the US on the ICRG "Law and Order" and the World Bank Institute considers only 5% of countries better in Rule of Law than Singapore⁷⁰. At the same time, it is classified as a dictatorship by the Cheibub and Gandhi classification. Table 6.4 shows the low correlation between the four Rule of Law indicators and the regime type variable from Cheibub and Gandhi.

Table 6.4 Correlations between Rule of Law indicators and regime type, income and growth				
	"Property Rights" Heritage Foundation	"Law and Order" ICRG	"Rule of Law" WGI	"Legal Structure and Security of Property Rights" Fraser Institute
Dictatorship Cheibub and Gandhi	-0.290	-0.150	-0.292	-0.267
Level of ln GDP per capita	0.717	0.838	0.681	0.735
Growth in ln GDP per capita	0.260	0.333	0.330	0.316

Table 6.5 lists the variables used, their definitions and sources. We control for the investment rate and trade openness with data from the World Development Indicators (2007) and for the initial income level with data from Heston, Summers and Aten (2006). We also

⁷⁰ Based on data for most recent year, obtained 3 August 2008 from the respective webpages, except for ICRG (data Dec year 2006).

6. Rule of Law and economic growth: A panel data re-examination

control for the quantity of human capital with the average annual growth in total population (World Development Indicators) and for the quality of human capital with average schooling years from Barro and Lee (2000). The availability of the Barro and Lee education variable severely limits our sample coverage. Unfortunately, alternative education variables from the World Development Indicators (labour force with primary, secondary and tertiary education) with a wider country coverage do not have sufficient historical depth.

Explanatory variables refer to the period of time (quinquennium) prior to the period of growth considered; in effect, they are lagged one period (initial income level, trade openness, regime type, education, health, investment rate, growth in population, Rule of Law).

Finally, we remark here that the four Rule of Law indicators are highly correlated with the level of GDP per capita and weakly correlated with growth in GDP per capita (see Table 6.4).

Table 6.5 Dependent variable and main control variables

Variable	Definition and Source
Growth in GDP per capita	Annual average growth rate over a period of five years. $(\ln \text{GDP per capita} - \text{five-year lag } \ln \text{gdp per capita}) * 100 / 5$
Ln GDP per capita	Ln of GDP per capita, Penn World Table 6.2, 2000 base year Source: Heston, Summers and Aten (2006)
Investment rate	Gross fixed capital formation (% of GDP) includes land improvements (fences, ditches, drains, and so on); plant, machinery, and equipment purchases; and the construction of roads, railways, and the like, including schools, offices, hospitals, private residential dwellings, and commercial and industrial buildings. According to the 1993 SNA, net acquisitions of valuables are also considered capital formation. Source: World Development Indicators (World Bank, 2007); World Bank national accounts data, and OECD National Accounts data files.
Population growth	Annual growth rate in total population. Total population is based on de facto definition of population, which counts all residents regardless of legal status or citizenship – except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. Source: World Development Indicators (World Bank, 2007); World Bank staff estimates from various sources including census reports, the United Nations Population Division's World Population Prospects, national statistical offices, household surveys conducted by national agencies, and Macro International.
Trade openness	Trade (% of GDP) is the sum of exports and imports of goods and services measured as a share of gross domestic product. Source: World Development Indicators (World Bank, 2007); World Bank national accounts data, and OECD National Accounts data files.
Dictatorship	coded 0 if democracy, 1 if dictatorship. A regime is considered a dictatorship if the chief executive is not elected, the legislature is not elected, there is no more than one party, or there has been no alternation in power (Przeworski <i>et al.</i> 2000). Transition years are coded as the regime that emerges in that year. Source: Cheibub and Gandhi (2004)
Education	Average schooling years in the total population aged 15 and over Source: Barro and Lee (2001); retrieved from QoG database.

6. Rule of Law and economic growth: A panel data re-examination

Health	Immunization, DPT (% of children ages 12-23 months) measures the percentage of children ages 12-23 months who received vaccinations before 12 months or at any time before the survey. A child is considered adequately immunized against DPT (diphtheria, pertussis or whooping cough, and tetanus) after receiving three doses of vaccine. Source: World Development Indicators (World Bank, 2007); World Health Organization and United Nations Children's Fund.
Temperate zones	Percentage of a country's population in Koeppen-Geiger temperate zone in 1995. Source: Center for International Development, Geography Data Sets (Gallup <i>et al.</i> , 1999)

6.3. Specification issues and estimation methods

Under classical but restrictive assumptions, Pooled OLS (POLS) will deliver consistent estimates of the coefficients and their standard errors. In particular, the panel members (countries) must be homogeneous in their parameters and all the observations must be independently sampled. Given the time series dimension in the data, however, it is clear that observations over time for a single country will not be independent; even in the absence of inertia and persistence, country heterogeneity (in the form of country effects) will create dependent blocks of observations. Fixed effects models deal with the heterogeneity by introducing arbitrary country effects and using only within-country variation to estimate the systematic effects of other variables. Random Effects models are a compromise: they allow for country effects but do not allow these to be arbitrary: they must be unrelated to any of the included explanatory variables. From the point of view of robustness, fixed effects estimates are the safest to use. Unfortunately, they are also the most demanding in terms of degrees of freedom, and provide no direct estimates of the effects of country-specific, time-constant variables.

The choice of method is by no means the only dilemma we face. We have discussed above Rule of Law indicators from four different sources, and although nothing in principle prevents one from including them collectively in growth regressions to evaluate their comparative effects and complementarities, it will soon become obvious that this is too much to be asked from the data. One obvious

6. Rule of Law and economic growth: A panel data re-examination

way to find out how well governance indicators stand the test of different definitions and metrics is to vary the source of the measurements and see whether comparable effects or patterns emerge.

The four sources have different coverage, both in terms of geography (countries) and in terms of history (years). This is the rationale for comparing them in two different ways: over the countries and years they have all four in common (“listwise deletion” of incomplete observations), and over the countries and years they have in common pairwise (“pairwise deletion” of incomplete observations).

Another choice to be made concerns the scaling of the governance indicators. For practical reasons and probably for convenience of communication, all four indicators are published in cardinal form, as numbers on a pseudo-metric scale (of 1 to 6 for ICRG; -2.5 to +2.5 for WGI; 0 to 10 for Fraser Institute; and 0 to 100 for the Heritage foundation ⁷¹). Fundamentally, however, the information they convey is at best of ordinal nature, and the scales are not always comparable over time, let alone across sources. Including such indicators with constant coefficients in linear regressions imposes disputable assumptions of cardinality and linearity. To relax these to some extent, we try to extract the main information in the data by a categorical transformation. We define “low” values of the governance indicators as values less than or equal to the lower quartile limit, and “high” values as those above the upper quartile limit. We call values in between the “median” values (the middle half of the

⁷¹ The Heritage Foundation Property Rights indicator was rescaled for our calculations, see Table 6.1.

sample or somewhat less). We then generate two binary indicator variables or dummies, one for the “low” and one for “high” categories. These transformations have advantages in terms of interpreting results: the magnitude of coefficients is easy to interpret and the estimated coefficients are directly comparable across indicators. The same advantages hold for interaction terms.

Next a selection has to be made among a very large set of potentially relevant control variables: We define a “long” list of regressors for the growth equations, that comprises the following: investment rate, population growth rate, initial income per capita, Rule of Law; dictatorship, education, health, trade openness and temperate zone. We consider pooled OLS, random effects and fixed effects specifications including all. To limit the proliferation of parameters and the loss of degrees of freedom, we also consider Pooled estimates and random effects estimates in which several control variables are replaced by time-constant averages or initial values⁷².

A final choice to be mentioned here concerns the treatment of the possible endogeneity of initial income. One concern is simultaneity between income and governance, although the time precedence of initial income level should alleviate this concern. Another concern is measurement error in income per capita levels, likely to cause on the one hand attenuation bias in its coefficient and on the other hand spurious correlation with growth (measured as the

⁷² Since few coefficients turned out significant, we also considered shortened lists of regressors and verified that the lack of significance was not due to multicollinearity.

6. Rule of Law and economic growth: A panel data re-examination

forward difference in income per capita). To deal with this possibility we run RE-IV and FE-IV regressions instrumenting initial income with its own lag. Table 6.6 summarises our selection of methods.

Table 6.6: Summary table of estimation methods

		List of regressors
P-OLS	Pooled OLS	Rule of Law; initial income per capita, investment rate, population growth rate, trade openness, dictatorship, education, health, temperate zones
RE1	Random Effects GLS with some variables replaced by time-averages	Time-varying: Rule of Law; initial income per capita, investment rate, population growth rate Time-averaged: trade openness, dictatorship, education, health, temperate zones
RE2	Random Effects GLS	Rule of Law; initial income per capita, investment rate, population growth rate, trade openness, dictatorship, education, health, temperate zones
FE	Fixed Effects OLS (=LSDV)	Rule of Law; initial income per capita, investment rate, population growth rate, trade openness, dictatorship, education, health, temperate zones
P-IV	Pooled OLS + IV treatment of income	
RE1-IV	RE1 with Time Averages of control variables and IV treatment of income	
RE2-IV	RE2 + IV treatment of income	
FE-IV	FE + IV treatment of income	

6.4 Empirical Results

6.4.1 Growth equations without governance indicators

As mentioned in Section 6.2 there is a large body of literature on the determinants of economic growth. In this subsection we estimate growth equations, at first excluding governance indicators. We regress growth in GDP per capita on five-year-lagged values of \ln GDP per capita, the investment rate and population growth, trade openness, dictatorship, education, health and the proportion of the population living in temperate zones. All specifications also include fixed time effects (by means of dummy variables, although to save space the corresponding coefficients are not reported in the tables). We vary the sample composition and the extent of heterogeneity allowed in order to assess how these affect the outcomes. We compare the following estimates, in increasing order of generality:

- Pooled OLS, treating the whole panel as if it constituted one large random sample (including fixed time effects)
- Random Effects GLS (RE1), replacing several control variables by their time averages (including fixed time effects)
- Random Effects GLS (RE2), including all control variables in their time-varying form (including fixed time effects)
- Fixed Effects (LSDV) estimation, allowing for both fixed time and fixed country effects.

Table 6.7(a) presents the estimates using the widest possible sample composition, Table 6.7(b) presents them for a smaller sample

6. Rule of Law and economic growth: A panel data re-examination

restricted to those countries and years for which all four Rule-of-Law indicators (to be discussed in the next subsection) are available. The first striking outcome is that the estimates are very sensitive to the sample composition. The coefficients of the two most classical determinants of growth, the investment rate and population growth, have varying signs and magnitudes and weak significance in the full samples; whereas in the restricted samples, the coefficient of the investment rate is close to 0.1 and significant for all methods used, and the coefficient of population growth is negative throughout (though insignificant under fixed effects estimation). Among the additional control variables, trade openness, education and temperate zone have some significant coefficients in the full samples, but none in the restricted one. The coefficient of the lagged income level is systematically negative, but of very variable magnitude and significance.

One possible problem with these estimates is that the lagged income level is measured with error and that (apart from attenuation bias) this induces spurious correlation and endogeneity with respect to growth (the lagged income level being the denominator of the growth variable). We therefore repeat the estimations of Table 7 treating the lagged income level as endogenous and instrumenting with the income level lagged a further five years. These estimates are presented in Table 8, with Panel (a) referring to the full samples and Panel (b) to the restricted samples. In spite of some changes the results are qualitatively very similar to those in Table 7, suggesting that measurement error in the income level is not causing major biases.

6.4.2 Introducing the Rule of Law indicators

In this subsection we extend the growth equations with the Rule of Law indicators described in Section 6.3, and vary the details of specification and estimation in several ways so as to investigate the robustness of the results. As in the preceding subsection, we consider different sample compositions (full samples/same samples) and estimation methods (depending on the treatment of heterogeneity and of the endogeneity of lagged income). In addition, we have four different sources of the Rule of Law indicator to consider, and two different ways of scaling the indicators (pseudo-metric/categorical). We test the fragility of results using the four different Rule of Law indicators interchangeably. All specifications also include fixed time effects (although to save space these are not reported in the tables).

Tables 6.9 to 6.11 present an overall view of the significance of effects of governance using different regression methods. (For the detailed estimation results, see Tables A.01-A.08 in the appendix.) Each table consists of one panel using the largest possible sample (panel (a)), and another panel using only the observations available for all the indicators (panel (b)). Reported standard errors are calculated with a heteroskedasticity and autocorrelation robust formula in Tables 6.9 and 6.10, but not in the IV estimations of Table 6.11.

We denote estimates not significantly different from 0 at the 10% level by 0; those significant at the 10% but not at the 5% level by their sign (-, +); those significant at 5% but not the 1% level by twice

6. Rule of Law and economic growth: A panel data re-examination

their sign (--, ++); those significant at the 1% level by three times their sign (---, +++).

Choice of indicator

In none of our regressions does the Heritage Foundation indicator have a significant effect⁷³. The ICRG indicator only shows significant effects with the “Low” governance quality dummy. The two remaining indicators, WGI and Fraser, more frequently (though not systematically) show significant effects.

Sample coverage

Since the four Rule of Law indicators are available for different countries and years, some differences might be due to differences in the dataset. The estimated coefficients and their statistical significance are indeed sensitive to the sample period used. In general, we find more significant effects under the “same samples” restriction than in the full samples, in spite of the diminished sample sizes. Nevertheless, the difference in significance between the different indicators cannot be explained by the choice of sample coverage. The contrast in significance between the Heritage and ICRG indicators on the one hand, and the WGI and Fraser on the other, remains in the common sample, at least using the pseudo-metric scaling (and in a modified form under the categorical scaling).

⁷³ This is in spite of the fact that, of the four indicators considered, we would argue it has the clearest conceptual definition, avoiding to contaminate Rule of Law aspects with other, prosperity-related aspects.

Scaling of indicator

Rule of Law indicators measured on a pseudo-metric scale from the Heritage Foundation and from ICRG fail to exhibit significant effects in the growth regressions. In none of the regressions using categorically rescaled indicators does the “High quality of governance” dummy variable have a significant positive effect on growth. The “Low quality of governance” dummy fares better, except with the Heritage Foundation indicator, where none of the Rule of Law indicators seems to matter. In these regressions, it seems that only “bad” governance matters, and only if it is measured by another source than the Heritage Foundation.

Choice of method

The effects are seen also to be sensitive to the treatment of heterogeneity. Using Pooled OLS and random effects the pseudo-metric indicators get significant effects, but with fixed effects these weaken, and they vanish altogether in the complete samples. In general, since the quality of governance is unlikely to change quickly over time and has only been “observed” relatively recently, we would expect its effects to be much harder to identify under a fixed effects specification (where all cross-sectional variation has been differenced away). Since the random effects with time averages is more economical with the variation in the data than the traditional random effects estimator, it also tends to produce statistically more significant effects. Our estimates conform with these expectations.

6. Rule of Law and economic growth: A panel data re-examination

Turning to the categorically scaled indicators, the “High” dummies have no significant coefficients and thus seems to be immaterial, whatever the method. The “Low” dummy has a negative coefficient in a number of cases, in pooled regressions, but also in random effects regressions in the case of the Fraser Institute. The only indicator that shows some significant effect when full allowance is made for heterogeneity (fixed effects models) is ICRG.

Instrumenting the lagged income level

The findings in term of significance of the indicators are unaffected when we treat the lagged income level as endogenous and instrument it with its own lag. If anything, the ICRG coefficients seem somewhat more significant, however we note that the standard errors are not strictly comparable since a formula robust to heteroskedasticity and serial correlation has been used to construct Table 10 but was not readily available in all cases covered by Table 6.11.

We conclude from these outcomes that none of the Rule of Law effects is very robust to specifying fixed effects, and few are robust to specifying random effects with individual-specific time averages. Pooled estimates are more significant but are in general biased (and in particular omitted variable bias may explain their apparent significance). Rescaling the indicators into categories so as to allow a degree of nonlinearity leads to asymmetric estimates of the effects of “high” and “low” rule-of-law quality: “low” measured quality of governance reduces growth, but “high” measured quality does not

significantly enhance it. Allowing for a certain degree of endogeneity of the income level does not affect these conclusions.

Nonlinearities

Various attempts to uncover stronger nonlinear relationship were not successful. In particular, we allowed for different effects of Rule of Law in groups of countries defined by their income levels or on a regional basis (World Bank classifications), to not avail.

6.4.3 Fragility of significance of other regressors

The significance of several of the other explanatory variables is highly sensitive to the choice of Rule of Law indicators.

Ln GDP per capita is significant in pooled cross-section when using the ICRG or the Fraser Institute indicator (Table 6.A.01(a)). This is due to the deeper historical depths, since the significance disappears in the more limited common dataset (Table 6.A.01(b)). It is highly significant in fixed-effects analysis when using any of the four Rule of Law indicators (see Table 6.A.03 and Table 6.A.04), for both the extended and the common dataset.

The investment rate tends to be highly significant and of the order of 0.1, at least under the common samples restriction (Table 6.A.01(b)). This is clearly the case in all pooled cross-section regressions and in the random effects models with time averages.

6. Rule of Law and economic growth: A panel data re-examination

However, the relationship breaks down in the ICRG and Fraser Institute cases when the samples are unrestricted and full heterogeneity is allowed for. Since the unrestricted samples involve many more years with these indicators the risk of structural instability is greater and this may cause the breakdown.

The coefficient of population growth is much more erratic. It is significantly negative and of the order of $-.2$ to $-.4$ in the pooled OLS estimates, and similarly in the random effects estimates under the common samples restriction; however, its statistical significance breaks down with fixed effects, and similarly with random effects in full samples.

The remaining five controls have largely insignificant coefficients, with some exceptions.

Trade openness is not significant in any of the pooled or random effects regressions, although its coefficient is almost always positive and of the order of 0.005 to 0.015 . It is significant only in fixed effects analysis (Table 6.A.04) when using the ICRG and the Fraser Institute indicator, although much less so when the dataset is reduced.

The coefficient of the dictatorship variable has varying signs and magnitudes and is hardly ever significant. Maybe this should not be surprising, given that there are few changes in regime type over time.

Education does not fare much better. Only in some random effects models with ICRG and Fraser Institute data does a significant positive impact on growth show up, of the order of .2 to .3. The other human capital variable, health, exhibits a similar pattern.

Finally, the geographical “temperate zone” variable, which is redundant in fixed effects models (since it is time-constant), always has a positive effect, though it is not usually significant. Its magnitude varies widely, from 0.2 to 1.8.

6.4.4 Rule of Law and the investment rate

In this subsection we investigate whether investments in physical capital might be themselves determined by the Rule of Law and therefore masking its effect to the extent these are indirect. Tables 6.B.01 to 6.B.04 report detailed estimates of an equation explaining the investment rate in every five-year period by income growth and the other main variables in the preceding five-year period. The predictor variables that matter most are the (lagged) investment ratio and population growth and, to a lesser extent, trade openness; however, few effects remain significant in the fixed effects model. In no single case does economic growth turn out to be a significant predictor of the investments rate. As to the Rule of Law, there is no persistent pattern and therefore (in our view) no robust effect. In the pooled and random effects RE1 models, a low Heritage Foundation score seems oddly enough to enhance investment, but this effect disappears in random effects RE2 and fixed effects models. In the fixed effects model and

6. Rule of Law and economic growth: A panel data re-examination

under the restriction of coinciding samples, a low Fraser Institute score seems to discourage investment. For the remainder, there seems to be no more significance than would be expected by chance. We conclude that there is no substantial indirect channel of causality from Rule of Law to growth through the investment ratio.

6.5 Conclusion

In this paper we compiled a comprehensive data base of available Rule of Law indicators and tested the robustness of the effects they are estimated to have on economic growth. We find little evidence of robust effects once heterogeneity across countries is allowed for. Among the four indicator sources considered, Rule of Law as measured by the Heritage foundation had the least, and by the Fraser Institute the most significant effects on growth. The pseudo-metric scalings commonly used for the indicators seem to have little virtue. In particular, the effects of “good” and “bad” governance, if present, are probably asymmetric. In terms of economic growth, having a good Rule of Law matters less (if at all) than not having a bad one.

On a cautionary note, what we find is a striking lack of confirmation of a systematic, measurable effect. We do not think that we have established that governance or the Rule of Law do not matter, but that synthetic indicators are likely to be either gross oversimplifications (not accounting for all sorts of specific

circumstances) or too vague in their underlying concepts⁷⁴. Therefore, great care should be exercised in drawing policy conclusions from synthetic indicators and basing fund allocation on them⁷⁵. Claims that improving the Rule of Law, as defined by popular indicators, should be the policy priority in all developing countries with supposedly weak Rule of Law systems are potentially overstated and cannot be sustained at present by econometric panel data analysis⁷⁶.

Several explanations are thinkable why other studies have found more significant effects. One is that they selected the “right” (most significant) indicator. Why studies based on one indicator rather than the other are more prominent may then be the consequence of ordinary publication bias. A second partial explanation is that many studies looked at levels rather than growth; see Rodrik (2008). A third reason is that many studies analysed a cross-sectional data set only and not a panel, therefore attributing country heterogeneity to the variation in whatever indicators were included. A fourth possible

⁷⁴ The indicators may also fail to take into account that while for property rights to be efficient they need to be secure, the opposite is not true. Secure property rights are not automatically efficient property rights (see North, 1990 on the efficiency of property rights).

⁷⁵ In particular, formal Rule of Law reforms in developing countries, according to blueprints from developed countries, that fail to take into account the informal institutions in place (see North, 1990 for an analysis of the importance of informal rules and their interaction with formal rules) may be at a high risk of failure and may do more harm than good.

⁷⁶ “It is not possible here to survey all the literature on what is in fact an extremely complex, multifaceted question about the relationship of the Rule of Law and economic development. The central point is that simplistic assertions such as have become common among aid agencies to the effect that ‘the Rule of Law’ *grosso modo* is necessary for development are at best badly oversimplified and probably misleading in many ways” (Carothers, 2003, p.7).

6. Rule of Law and economic growth: A panel data re-examination

reason is that many studies omitted a proper human capital measure; see Glaeser *et al.* (2004). We think further research is warranted on these issues and meanwhile empirical analysts should be careful not to overstate the case and policy makers should be careful not to weigh weak evidence too heavily.

Tables 6.7 to 6.12

Table 6.7(a): Different methods, full samples, ordinary SEs

VARIABLES	(1) Pooled OLS	(2) RE1	(3) RE	(4) FE
ln GDP per cap.	-0.474* (0.255)	-1.067*** (0.308)	-1.259*** (0.338)	-7.895*** (0.685)
Investment rate	0.0839*** (0.0225)	0.0672*** (0.0250)	0.0405 (0.0270)	-0.0196 (0.0322)
Population growth	-0.216* (0.119)	0.147 (0.127)	0.0108 (0.133)	0.195 (0.142)
Trade openness	0.00210 (0.00401)	-0.00148 (0.00631)	0.00819 (0.00529)	0.0325*** (0.00843)
Dictatorship	-0.351 (0.348)	-1.040 (0.672)	-0.477 (0.396)	-0.154 (0.453)
Education	0.123 (0.0961)	0.209 (0.146)	0.338*** (0.130)	0.196 (0.290)
Health	0.0124 (0.00895)	0.0319** (0.0152)	0.0131 (0.0100)	0.00622 (0.0108)
Temperate zone	0.815* (0.485)	1.351** (0.679)	1.668** (0.688)	
Observations	358	373	358	379
Number of countries		95	94	100
R-squared	0.699	.	.	0.420
RMS Residual	2.410	2.081	2.017	1.691

6. Rule of Law and economic growth: A panel data re-examination

Table 6.7(b): Different methods, same samples, ordinary SEs

VARIABLES	(1) Pooled OLS	(2) RE1	(3) RE	(4) FE
ln GDP per cap.	-0.0935 (0.284)	-0.350 (0.411)	-0.465 (0.396)	-13.13*** (1.525)
Investment rate	0.106*** (0.0294)	0.110*** (0.0340)	0.104*** (0.0335)	0.0926** (0.0425)
Population growth	-0.399*** (0.129)	-0.287* (0.152)	-0.443*** (0.149)	-0.0203 (0.210)
Trade openness	0.00514 (0.00436)	0.00961 (0.00849)	0.00590 (0.00559)	0.0149 (0.0114)
Dictatorship	0.181 (0.393)	-0.101 (0.818)	0.386 (0.493)	0.254 (0.778)
Education	-0.0233 (0.103)	0.0791 (0.178)	0.0835 (0.147)	0.428 (0.798)
Health	0.00874 (0.0118)	0.00620 (0.0205)	0.00862 (0.0151)	0.00713 (0.0256)
Temperate zone	0.683 (0.543)	0.768 (0.867)	0.887 (0.773)	
Observations	160	163	160	165
Number of countries		83	82	85
RMS Residual	1.799	1.189	1.173	0.970
R-squared	0.777	.	.	0.613

Regressand leads regressors by one five-year period; fixed time effects included.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.8(a): Instrumenting income, full samples

VARIABLES	(1) Pooled IV	(2) RE1 IV	(3) RE2 IV	(4) FE IV
ln GDP per cap.	-0.885*** (0.265)	-1.406*** (0.322)	-1.596*** (0.360)	-9.226*** (0.887)
Investment rate	0.0847*** (0.0225)	0.0714*** (0.0250)	0.0416 (0.0271)	-0.00802 (0.0328)
Population growth	-0.215* (0.120)	0.158 (0.127)	0.00948 (0.133)	0.164 (0.143)
Trade openness	0.00209 (0.00403)	-0.00125 (0.00629)	0.00831 (0.00532)	0.0334*** (0.00849)
Dictatorship	-0.484 (0.350)	-1.214* (0.671)	-0.546 (0.398)	-0.129 (0.456)
Education	0.217** (0.0976)	0.278* (0.146)	0.418*** (0.134)	0.218 (0.293)
Health	0.0159* (0.00900)	0.0367** (0.0152)	0.0149 (0.0101)	0.00684 (0.0108)
Temperate zone	0.962** (0.488)	1.439** (0.677)	1.815*** (0.695)	
Observations	358	373	358	379
RMS Residual	2.419			
Number of countries		95	94	100

Regressand leads regressors by one five-year period; fixed time effects included.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.8(b): Instrumenting income, same samples

VARIABLES	(1) Pooled IV	(2) RE1 IV	(3) RE2 IV	(4) FE IV
ln GDP per cap.	-0.428 (0.291)	-0.464 (0.420)	-0.628 (0.412)	-11.95*** (2.683)
Investment rate	0.107*** (0.0296)	0.113*** (0.0337)	0.104*** (0.0334)	0.0908** (0.0428)
Population growth	-0.406*** (0.129)	-0.277* (0.151)	-0.444*** (0.149)	-0.0494 (0.218)
Trade openness	0.00524 (0.00438)	0.00959 (0.00828)	0.00593 (0.00557)	0.0132 (0.0119)
Dictatorship	0.122 (0.395)	-0.171 (0.801)	0.355 (0.492)	0.352 (0.802)
Education	0.0423 (0.105)	0.0966 (0.174)	0.116 (0.148)	0.343 (0.817)
Health	0.0132 (0.0119)	0.00809 (0.0201)	0.0106 (0.0151)	0.00383 (0.0264)
Temperate zone	0.869 (0.547)	0.830 (0.847)	0.984 (0.772)	
Observations	160	163	160	165
Number of countries		83	82	85
RMS Residual	1.808			

Regressand leads regressors by one five-year period; fixed time effects included.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.9: Pseudo-metric Rule of Law indicators

Table 6.9: Pseudo-metric Rule of Law indicators					
		Heritage	WGI	ICRG	Fraser
Panel 6.9(a): Full samples					
Pooled OLS		0	++	0	++
RE1		0	+++	0	+
RE2		0	++	0	0
FE		0	0	0	0
Panel 6.9(b): Same samples					
Pooled OLS		0	+++	0	+++
RE1		0	+++	0	+++
RE2		0	++	0	++
FE		0	0	0	+

Notes to Table 6.9: Robust standard errors used

0 = not significantly different from 0 at the 10% level

-, --, --- = significantly negative at the 10%, 5%, 1% level, respectively

+, ++, +++ = significantly positive at the 10%, 5%, 1% level, respectively.

6. Rule of Law and economic growth: A panel data re-examination

Table 6.10: Categorical Rule of Law dummies

Table 6.10: Categorical Rule of Law dummies									
		Heritage		WGI		ICRG		Fraser	
		Low	High	Low	High	Low	High	Low	High
Panel 6.10(a): Full samples									
Pooled OLS		0	0	--	0	--	0	---	0
RE1		0	0	0	0	0	0	--	0
RE2		0	0	0	0	0	0	0	0
FE		0	0	0	-	0	0	0	0
Panel 6.10(b): Same samples									
Pooled OLS		0	0	---	0	-	0	---	0
RE1		0	0	-	0	-	0	--	0
RE2		0	0	0	0	0	0	--	0
FE		0	0	0	0	-	0	0	0

Notes to Table 6.10: Robust standard errors used

0 = not significantly different from 0 at the 10% level

-, --, --- = significantly negative at the 10%, 5%, 1% level, respectively

+, ++, +++ = significantly positive at the 10%, 5%, 1% level, respectively.

Table 6.11: Categorical Rule of Law dummies, instrumenting income

Table 6.11: Categorical Rule of Law dummies, instrumenting income									
		Heritage		WGI		ICRG		Fraser	
		Low	High	Low	High	Low	High	Low	High
Panel 6.11(a): Full samples									
Pooled IV		0	0	--	0	--	0	---	0
RE1-IV		0	0	0	0	-	0	--	0
RE2-IV		0	0	0	0	0	0	0	0
FE-IV		0	0	0	0	0	0	0	0
Panel 6.11(b): Same samples									
Pooled IV		0	0	---	0	-	0	---	0
RE1-IV		0	0	-	0	--	0	--	0
RE2-IV		0	0	0	0	-	0	--	0
FE-IV		0	0	0	0	--	0	0	0

Notes to Table 6.11: Traditional (non robust) standard errors used for RE1-IV, RE2-IV, and FE methods
0 = not significantly different from 0 at the 10% level
-, --, --- = significantly negative at the 10%, 5%, 1% level, respectively
+, ++, +++ = significantly positive at the 10%, 5%, 1% level, respectively.

6. Rule of Law and economic growth: A panel data re-examination

Table 6.12: Categorical Rule of Law dummies in the investment equation									
		Heritage		WGI		ICRG		Fraser	
		Low	High	Low	High	Low	High	Low	High
Panel 6.12(a): Full samples									
Pooled OLS		+++	0	+	0	0	0	0	0
RE1		++	0	0	0	0	0	0	0
RE2		0	0	0	0	-	0	0	0
FE		0	0	+	0	0	0	0	0
Panel 6.12(b): Same samples									
Pooled OLS		+++	0	0	0	0	0	0	0
RE1		++	0	0	0	0	0	0	0
RE2		0	0	0	0	0	0	0	0
FE		0	0	0	0	0	0	--	0

Notes to Table 6.12: Robust standard errors used
0 = not significantly different from 0 at the 10% level
-, --, --- = significantly negative at the 10%, 5%, 1% level, respectively
+, ++, +++ = significantly positive at the 10%, 5%, 1% level, respectively.

APPENDIX: Detailed Tables

Table 6.A.01(a): Pooled OLS, full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	0.0710 (0.528)	-1.035** (0.453)	-0.834** (0.385)	-0.958*** (0.356)
Rule of Law high	-0.168 (0.413)	0.242 (0.509)	-0.650 (0.451)	0.198 (0.588)
ln GDP per cap.	0.165 (0.329)	-0.0269 (0.338)	-0.681** (0.329)	-0.716** (0.337)
Investment rate	0.0777** (0.0343)	0.0728** (0.0325)	0.0952*** (0.0358)	0.107*** (0.0363)
Population growth	-0.278** (0.120)	-0.234** (0.118)	-0.333*** (0.127)	-0.218* (0.115)
Trade openness	0.00464 (0.00471)	0.00418 (0.00462)	0.00145 (0.00429)	0.00102 (0.00434)
Dictatorship	0.274 (0.421)	0.420 (0.431)	-0.471 (0.381)	-0.410 (0.355)
Education	-0.0405 (0.0974)	-0.0356 (0.0947)	0.112 (0.0924)	0.107 (0.0927)
Health	0.00509 (0.0148)	-0.00115 (0.0141)	0.0143 (0.0104)	0.0152 (0.0104)
Temperate zone	0.640 (0.576)	0.627 (0.557)	0.975* (0.546)	0.700 (0.576)
Observations	177	180	332	332
R-squared	0.763	0.769	0.708	0.713
RMS Residual	1.877	1.842	2.400	2.363

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.01(b): Pooled OLS, same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-0.438 (0.557)	-1.403*** (0.468)	-0.731* (0.425)	-1.328*** (0.361)
Rule of Law high	-0.447 (0.454)	0.198 (0.509)	-0.265 (0.477)	0.349 (0.484)
ln GDP per cap.	-0.0231 (0.331)	-0.253 (0.353)	-0.0863 (0.330)	-0.168 (0.331)
Investment rate	0.110*** (0.0420)	0.102** (0.0413)	0.0938** (0.0445)	0.105** (0.0425)
Population growth	-0.428*** (0.131)	-0.381*** (0.129)	-0.406*** (0.131)	-0.318** (0.133)
Trade openness	0.00526 (0.00463)	0.00540 (0.00445)	0.00558 (0.00462)	0.00281 (0.00469)
Dictatorship	0.298 (0.452)	0.337 (0.454)	0.0887 (0.474)	0.277 (0.420)
Education	-0.0205 (0.0879)	-0.0319 (0.0870)	-0.0548 (0.0908)	-0.0215 (0.0868)
Health	0.00437 (0.0149)	0.000794 (0.0135)	0.00519 (0.0146)	0.00256 (0.0138)
Temperate zone	0.857 (0.583)	0.739 (0.555)	0.696 (0.576)	0.255 (0.574)
Observations	160	160	160	160
R-squared	0.780	0.789	0.782	0.794
RMS Residual	1.802	1.763	1.791	1.741

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6.A.02(a): RE1 (Time Averages of control variables), full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-0.206 (0.566)	-0.461 (0.451)	-0.603 (0.404)	-0.729** (0.350)
Rule of Law high	-0.154 (0.392)	0.249 (0.507)	-0.466 (0.477)	0.290 (0.624)
ln GDP per cap.	-0.139 (0.440)	-0.231 (0.411)	-1.192*** (0.410)	-1.253*** (0.419)
Investment rate	0.0850** (0.0394)	0.0802** (0.0365)	0.0784** (0.0371)	0.0744** (0.0374)
Population growth	-0.0770 (0.136)	-0.0544 (0.133)	0.0772 (0.147)	0.130 (0.126)
TA _v trade openness	0.00644 (0.00798)	0.00669 (0.00760)	-0.00342 (0.00753)	-0.00290 (0.00732)
TA _v dictatorship	0.126 (0.787)	0.127 (0.808)	-1.292* (0.709)	-1.114 (0.688)
TA _v education	0.111 (0.157)	0.102 (0.150)	0.190 (0.130)	0.169 (0.128)
TA _v health	-0.00245 (0.0227)	-0.00430 (0.0214)	0.0349* (0.0185)	0.0425** (0.0182)
Temperate zone	0.854 (0.767)	0.789 (0.762)	1.369* (0.734)	1.034 (0.766)
Observations	184	187	341	344
Number of countries	94	95	87	88
RMS Residual	1.233	1.240	2.120	2.078

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.02(b): RE1 (Time Averages of control variables), same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-0.704 (0.649)	-0.827* (0.486)	-0.985* (0.525)	-0.900** (0.417)
Rule of Law high	-0.563 (0.393)	0.240 (0.517)	0.208 (0.420)	0.581 (0.448)
ln GDP per cap.	-0.304 (0.499)	-0.392 (0.466)	-0.452 (0.473)	-0.337 (0.442)
Investment rate	0.109** (0.0497)	0.105** (0.0491)	0.102** (0.0508)	0.111** (0.0488)
Population growth	-0.307** (0.146)	-0.288* (0.147)	-0.294* (0.151)	-0.240 (0.151)
TA _v trade openness	0.0103 (0.00890)	0.00982 (0.00820)	0.0104 (0.00867)	0.00702 (0.00837)
TA _v dictatorship	0.143 (0.819)	0.0903 (0.821)	-0.254 (0.874)	0.0537 (0.775)
TA _v education	0.0996 (0.157)	0.0643 (0.147)	0.0372 (0.156)	0.0599 (0.141)
TA _v health	0.00111 (0.0256)	0.000115 (0.0230)	-0.00119 (0.0245)	-0.00188 (0.0230)
Temperate zone	1.053 (0.831)	0.748 (0.816)	0.602 (0.842)	0.341 (0.800)
Observations	163	163	163	163
RMS Residual	1.174	1.214	1.175	1.212
Number of countries	83	83	83	83

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.A.03(a): RE2 full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-0.153 (0.599)	-0.510 (0.446)	-0.508 (0.406)	-0.555 (0.347)
Rule of Law high	-0.188 (0.381)	0.218 (0.504)	-0.393 (0.465)	0.386 (0.617)
ln GDP per cap.	-0.181 (0.491)	-0.335 (0.476)	-1.379*** (0.440)	-1.328*** (0.434)
Investment rate	0.0795* (0.0413)	0.0718* (0.0385)	0.0460 (0.0388)	0.0470 (0.0393)
Population growth	-0.154 (0.152)	-0.121 (0.147)	-0.0814 (0.147)	-0.0216 (0.123)
Trade openness	0.00434 (0.00527)	0.00464 (0.00522)	0.00737 (0.00553)	0.00739 (0.00552)
Dictatorship	0.311 (0.539)	0.308 (0.538)	-0.542 (0.444)	-0.492 (0.404)
Education	0.0606 (0.147)	0.0613 (0.140)	0.293** (0.128)	0.275** (0.124)
Health	0.00428 (0.0190)	0.00420 (0.0181)	0.0184* (0.0112)	0.0179 (0.0111)
Temperate zone	0.997 (0.738)	0.930 (0.737)	1.809** (0.741)	1.431* (0.755)
Observations	177	180	332	332
RMS Residual	1.241	1.237	2.035	2.025
Number of countries	91	92	86	87
R-squared

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.03(b): RE2 same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-0.698 (0.650)	-0.745 (0.473)	-0.785 (0.537)	-0.819** (0.406)
Rule of Law high	-0.589 (0.396)	0.146 (0.510)	0.214 (0.434)	0.638 (0.442)
ln GDP per cap.	-0.416 (0.517)	-0.547 (0.509)	-0.565 (0.501)	-0.539 (0.474)
Investment rate	0.108** (0.0507)	0.0985* (0.0506)	0.0978* (0.0535)	0.104** (0.0504)
Population growth	-0.463*** (0.152)	-0.429*** (0.148)	-0.423*** (0.155)	-0.384** (0.150)
Trade openness	0.00572 (0.00510)	0.00610 (0.00493)	0.00558 (0.00505)	0.00461 (0.00505)
Dictatorship	0.506 (0.582)	0.396 (0.576)	0.204 (0.631)	0.429 (0.547)
Education	0.0904 (0.140)	0.0690 (0.132)	0.0508 (0.137)	0.0570 (0.130)
Health	0.00235 (0.0201)	0.00593 (0.0185)	0.00610 (0.0194)	0.00711 (0.0187)
Temperate zone	1.174 (0.765)	0.870 (0.760)	0.674 (0.791)	0.426 (0.744)
Observations	160	160	160	160
R-squared
RMS Residual	1.155	1.189	1.159	1.183
Number of countries	82	82	82	82

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.A.04(a): FE full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-0.721 (0.868)	-0.689 (0.489)	-0.0215 (0.419)	-0.372 (0.315)
Rule of Law high	-0.352 (0.555)	-0.758* (0.415)	-0.365 (0.428)	-0.0813 (0.499)
ln GDP per cap.	-12.51*** (1.692)	-12.45*** (1.728)	-7.693*** (1.151)	-7.621*** (1.140)
Investment rate	0.0682* (0.0396)	0.0544 (0.0388)	-0.0474 (0.0359)	-0.0674** (0.0332)
Population growth	0.135 (0.113)	0.196* (0.113)	0.204 (0.212)	0.198 (0.153)
Trade openness	0.0145 (0.0124)	0.0142 (0.0117)	0.0348*** (0.0122)	0.0340*** (0.0119)
Dictatorship	0.00925 (0.518)	0.00694 (0.569)	-0.119 (0.516)	-0.131 (0.484)
Education	-0.119 (0.822)	0.287 (0.790)	0.187 (0.297)	0.140 (0.273)
Health	-0.00189 (0.0206)	0.0125 (0.0210)	0.0142 (0.0121)	0.00850 (0.0116)
Observations	187	191	341	351
Number of countries	97	98	89	93
R-squared	0.561	0.554	0.423	0.444
RMS Residual	0.702	0.711	1.458	1.398

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.04(b): FE same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law low	-1.028 (0.973)	-0.428 (0.454)	-1.514* (0.890)	-0.465 (0.517)
Rule of Law high	-0.555 (0.567)	-0.643 (0.467)	0.217 (0.340)	0.506 (0.393)
ln GDP per cap.	-12.97*** (1.835)	-13.28*** (1.838)	-13.78*** (1.845)	-13.21*** (1.842)
Investment rate	0.0846* (0.0431)	0.0879** (0.0439)	0.0860** (0.0411)	0.0925** (0.0445)
Population growth	0.00390 (0.201)	0.0255 (0.186)	-0.0378 (0.176)	-0.0207 (0.197)
Trade openness	0.0126 (0.0103)	0.0129 (0.0106)	0.0171* (0.00950)	0.0168* (0.00995)
Dictatorship	0.133 (0.524)	0.130 (0.544)	-0.587 (0.686)	0.307 (0.460)
Education	0.265 (0.811)	0.452 (0.772)	0.535 (0.855)	0.332 (0.776)
Health	-0.000544 (0.0181)	0.00794 (0.0202)	0.0127 (0.0192)	0.0115 (0.0211)
Observations	165	165	165	165
Number of countries	85	85	85	85
R-squared	0.636	0.623	0.644	0.624
RMS Residual	0.643	0.655	0.637	0.654

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.A.05(a): Pooled IV, full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-0.232 (0.318)	-0.426 (0.327)	-1.100*** (0.360)	-1.101*** (0.371)
Rule of Law low	0.0417 (0.531)	-1.133** (0.463)	-0.859** (0.385)	-0.960*** (0.358)
Rule of Law high	0.0360 (0.422)	0.510 (0.497)	-0.527 (0.459)	0.341 (0.583)
Investment rate	0.0759** (0.0353)	0.0725** (0.0335)	0.102*** (0.0360)	0.113*** (0.0365)
Population growth	-0.275** (0.121)	-0.232* (0.120)	-0.327** (0.128)	-0.208* (0.117)
Trade openness	0.00451 (0.00470)	0.00410 (0.00460)	0.00136 (0.00425)	0.00113 (0.00428)
Dictatorship	0.217 (0.424)	0.373 (0.436)	-0.607 (0.389)	-0.531 (0.360)
Education	0.0400 (0.0931)	0.0407 (0.0908)	0.191** (0.0949)	0.177* (0.0948)
Health	0.00968 (0.0154)	0.00334 (0.0146)	0.0172 (0.0105)	0.0183* (0.0104)
Temperate zone	0.732 (0.583)	0.677 (0.562)	1.124** (0.552)	0.831 (0.582)
Observations	177	180	332	332
RMS Residual	1.886	1.852	2.408	2.371

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.05(b): Pooled IV, same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-0.398 (0.327)	-0.618* (0.347)	-0.426 (0.315)	-0.489 (0.316)
Rule of Law low	-0.481 (0.560)	-1.501*** (0.480)	-0.736* (0.426)	-1.338*** (0.364)
Rule of Law high	-0.245 (0.458)	0.456 (0.499)	-0.120 (0.475)	0.519 (0.476)
Investment rate	0.111*** (0.0425)	0.104** (0.0413)	0.0969** (0.0447)	0.108** (0.0424)
Population growth	-0.428*** (0.130)	-0.380*** (0.129)	-0.404*** (0.131)	-0.314** (0.134)
Trade openness	0.00516 (0.00458)	0.00526 (0.00438)	0.00547 (0.00454)	0.00267 (0.00459)
Dictatorship	0.245 (0.453)	0.289 (0.456)	0.0229 (0.477)	0.221 (0.419)
Education	0.0408 (0.0846)	0.0225 (0.0840)	0.00661 (0.0857)	0.0330 (0.0830)
Health	0.00903 (0.0154)	0.00507 (0.0139)	0.00953 (0.0152)	0.00680 (0.0144)
Temperate zone	0.977* (0.589)	0.818 (0.559)	0.806 (0.583)	0.352 (0.580)
Observations	160	160	160	160
RMS Residual	1.811	1.771	1.799	1.748

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.A.06(a): RE-IV-TAv (Time Averages of control variables), full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-0.367 (0.401)	-0.456 (0.389)	-1.549*** (0.342)	-1.600*** (0.337)
Rule of Law low	-0.229 (0.502)	-0.534 (0.436)	-0.648* (0.370)	-0.758** (0.331)
Rule of Law high	-0.0817 (0.439)	0.377 (0.596)	-0.430 (0.457)	0.353 (0.531)
Investment rate	0.0862*** (0.0279)	0.0822*** (0.0269)	0.0870*** (0.0298)	0.0845*** (0.0298)
Population growth	-0.0688 (0.121)	-0.0459 (0.118)	0.0861 (0.143)	0.141 (0.132)
TAv trade openness	0.00650 (0.00747)	0.00660 (0.00710)	-0.00313 (0.00684)	-0.00231 (0.00665)
TAv dictatorship	0.0259 (0.804)	0.0395 (0.752)	-1.480** (0.682)	-1.299** (0.657)
TAv education	0.153 (0.171)	0.140 (0.165)	0.247* (0.145)	0.221 (0.143)
TAv health	0.000754 (0.0184)	-0.00142 (0.0175)	0.0390** (0.0160)	0.0468*** (0.0156)
Temperate zone	0.903 (0.788)	0.811 (0.773)	1.491** (0.732)	1.154 (0.730)
Observations	184	187	341	344
Number of countries	94	95	87	88

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.06(b): RE-IV-TAv (Time Averages of control variables), same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-0.404 (0.445)	-0.555 (0.431)	-0.559 (0.423)	-0.467 (0.397)
Rule of Law low	-0.715 (0.546)	-0.885* (0.485)	-0.995** (0.489)	-0.950** (0.413)
Rule of Law high	-0.525 (0.461)	0.329 (0.608)	0.217 (0.509)	0.599 (0.569)
Investment rate	0.111*** (0.0341)	0.107*** (0.0335)	0.105*** (0.0354)	0.114*** (0.0333)
Population growth	-0.300** (0.153)	-0.277* (0.148)	-0.284* (0.152)	-0.226 (0.148)
TAv trade openness	0.0102 (0.00849)	0.00976 (0.00805)	0.0103 (0.00834)	0.00688 (0.00783)
TAv dictatorship	0.0941 (0.844)	0.0229 (0.777)	-0.321 (0.803)	-0.0192 (0.751)
TAv education	0.113 (0.178)	0.0848 (0.168)	0.0527 (0.175)	0.0798 (0.164)
TAv health	0.00261 (0.0210)	0.00234 (0.0195)	0.000491 (0.0204)	-6.20e-05 (0.0189)
Temperate zone	1.087 (0.884)	0.791 (0.851)	0.655 (0.875)	0.393 (0.825)
Observations	163	163	163	163
Number of countries	83	83	83	83

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.A.07(a): RE2-IV full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-0.487 (0.419)	-0.636 (0.413)	-1.726*** (0.382)	-1.719*** (0.367)
Rule of Law low	-0.179 (0.500)	-0.559 (0.446)	-0.519 (0.371)	-0.543 (0.337)
Rule of Law high	-0.107 (0.433)	0.354 (0.591)	-0.333 (0.449)	0.472 (0.526)
Investment rate	0.0783*** (0.0289)	0.0715** (0.0280)	0.0512* (0.0311)	0.0498 (0.0311)
Population growth	-0.155 (0.121)	-0.121 (0.119)	-0.0861 (0.151)	-0.0156 (0.138)
Trade openness	0.00433 (0.00533)	0.00465 (0.00521)	0.00736 (0.00568)	0.00785 (0.00558)
Dictatorship	0.277 (0.480)	0.268 (0.459)	-0.616 (0.424)	-0.565 (0.403)
Education	0.129 (0.148)	0.125 (0.145)	0.363*** (0.134)	0.355*** (0.130)
Health	0.00737 (0.0150)	0.00729 (0.0145)	0.0201* (0.0107)	0.0199* (0.0104)
Temperate zone	1.105 (0.739)	1.001 (0.739)	1.966*** (0.747)	1.627** (0.744)
Observations	177	180	332	332
Number of countries	91	92	86	87

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.07(b): RE2-IV same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-0.561 (0.439)	-0.743* (0.433)	-0.730* (0.421)	-0.715* (0.402)
Rule of Law low	-0.716 (0.520)	-0.773* (0.466)	-0.796* (0.474)	-0.833** (0.396)
Rule of Law high	-0.552 (0.447)	0.232 (0.592)	0.238 (0.491)	0.685 (0.541)
Investment rate	0.108*** (0.0338)	0.0989*** (0.0334)	0.0990*** (0.0355)	0.105*** (0.0332)
Population growth	-0.464*** (0.151)	-0.430*** (0.148)	-0.422*** (0.152)	-0.383*** (0.148)
Trade openness	0.00575 (0.00565)	0.00613 (0.00550)	0.00559 (0.00563)	0.00459 (0.00541)
Dictatorship	0.491 (0.505)	0.372 (0.487)	0.173 (0.506)	0.398 (0.478)
Education	0.118 (0.152)	0.103 (0.146)	0.0823 (0.152)	0.0897 (0.142)
Health	0.00385 (0.0157)	0.00809 (0.0152)	0.00797 (0.0152)	0.00923 (0.0147)
Temperate zone	1.245 (0.807)	0.943 (0.795)	0.758 (0.815)	0.504 (0.773)
Observations	160	160	160	160
Number of countries	82	82	82	82

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.A.08(a): FE-IV full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-14.38*** (3.086)	-13.40*** (3.093)	-9.147*** (0.929)	-9.002*** (0.873)
Rule of Law low	-0.762 (0.645)	-0.751 (0.518)	-0.0343 (0.363)	-0.428 (0.315)
Rule of Law high	-0.304 (0.417)	-0.732 (0.631)	-0.300 (0.423)	-0.101 (0.506)
Investment rate	0.0706* (0.0413)	0.0547 (0.0406)	-0.0333 (0.0362)	-0.0523 (0.0358)
Population growth	0.140 (0.138)	0.205 (0.146)	0.153 (0.178)	0.158 (0.149)
Trade openness	0.0177 (0.0125)	0.0159 (0.0127)	0.0353*** (0.00908)	0.0346*** (0.00859)
Dictatorship	-0.141 (0.861)	-0.0881 (0.883)	-0.0977 (0.503)	-0.0641 (0.469)
Education	-0.0110 (0.863)	0.360 (0.876)	0.211 (0.301)	0.147 (0.291)
Health	0.00331 (0.0266)	0.0153 (0.0256)	0.0150 (0.0116)	0.00853 (0.0109)
Observations	187	191	341	351
Number of countries	97	98	89	93

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.A.08(b): FE-IV same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
ln GDP per cap.	-11.70*** (2.683)	-12.27*** (2.774)	-12.92*** (2.625)	-12.09*** (2.690)
Rule of Law low	-0.998 (0.635)	-0.377 (0.495)	-1.445** (0.672)	-0.434 (0.431)
Rule of Law high	-0.608 (0.438)	-0.665 (0.646)	0.224 (0.464)	0.534 (0.577)
Investment rate	0.0833* (0.0426)	0.0869** (0.0430)	0.0855* (0.0445)	0.0911** (0.0429)
Population growth	-0.0227 (0.215)	0.00120 (0.221)	-0.0573 (0.211)	-0.0485 (0.218)
Trade openness	0.0105 (0.0120)	0.0113 (0.0124)	0.0158 (0.0116)	0.0151 (0.0120)
Dictatorship	0.235 (0.791)	0.226 (0.823)	-0.480 (0.881)	0.396 (0.801)
Education	0.173 (0.808)	0.365 (0.840)	0.471 (0.796)	0.249 (0.822)
Health	-0.00401 (0.0263)	0.00497 (0.0266)	0.0101 (0.0259)	0.00815 (0.0267)
Observations	165	165	165	165
Number of countries	85	85	85	85

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.B.01(a): Investment rate, Pooled OLS, full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	2.066*** (0.663)	1.252* (0.642)	-0.343 (0.630)	-0.123 (0.565)
Rule of Law >Q3	-0.569 (0.687)	0.760 (0.712)	-0.432 (0.603)	-0.355 (0.666)
gy62	-0.0282 (0.150)	-0.0130 (0.127)	-0.0176 (0.105)	-0.0367 (0.115)
Investment rate	0.659*** (0.0409)	0.677*** (0.0402)	0.667*** (0.0515)	0.689*** (0.0554)
Population growth	-0.415** (0.195)	-0.456** (0.180)	-0.343 (0.224)	-0.256 (0.210)
Trade openness	-0.0109* (0.00559)	-0.0136** (0.00585)	-0.00560 (0.00757)	-0.00582 (0.00781)
Dictatorship	-0.273 (0.585)	0.174 (0.630)	0.0538 (0.554)	-0.178 (0.536)
Education	-0.0821 (0.115)	-0.178 (0.126)	-0.150 (0.109)	-0.0865 (0.115)
Health	0.0153 (0.0210)	0.00666 (0.0207)	0.0245* (0.0135)	0.0228* (0.0134)
Temperate zone	0.780 (0.966)	0.153 (0.824)	-0.409 (0.784)	-0.369 (0.805)
Observations	177	180	332	332
R-squared	0.983	0.982	0.971	0.972
RMS Residual	2.898	2.932	3.685	3.631

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.B.01(b): Investment rate, Pooled OLS, same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	2.286*** (0.748)	0.846 (0.795)	-0.230 (0.681)	-0.0369 (0.663)
Rule of Law >Q3	-0.551 (0.761)	0.776 (0.747)	0.975 (0.889)	0.964 (0.755)
gy62	-0.0249 (0.168)	-0.0510 (0.176)	-0.0571 (0.178)	-0.0685 (0.175)
Investment rate	0.653*** (0.0638)	0.677*** (0.0681)	0.685*** (0.0670)	0.688*** (0.0678)
Population growth	-0.567*** (0.210)	-0.592*** (0.222)	-0.540** (0.240)	-0.557** (0.225)
Trade openness	-0.0115* (0.00645)	-0.0141** (0.00675)	-0.0145** (0.00682)	-0.0143** (0.00682)
Dictatorship	-0.270 (0.661)	0.199 (0.737)	0.221 (0.764)	0.283 (0.759)
Education	-0.0730 (0.129)	-0.180 (0.137)	-0.198 (0.133)	-0.202 (0.135)
Health	0.0160 (0.0222)	0.00462 (0.0226)	-0.00451 (0.0223)	-0.00185 (0.0222)
Temperate zone	0.468 (1.041)	-0.127 (0.902)	-0.257 (1.002)	-0.152 (0.938)
Observations	160	160	160	160
R-squared	0.982	0.981	0.981	0.981
RMS Residual	2.980	3.061	3.066	3.068

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.B.02(a): Investment rate, RE full samples with time averages of control variables

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	1.846** (0.775)	0.576 (0.780)	-0.111 (0.598)	-0.156 (0.548)
Rule of Law >Q3	-0.560 (0.734)	1.100 (0.804)	-0.158 (0.609)	-0.280 (0.666)
gy62	0.0311 (0.158)	0.0158 (0.133)	0.00351 (0.100)	-0.0231 (0.109)
Investment rate	0.632*** (0.0449)	0.632*** (0.0434)	0.672*** (0.0499)	0.701*** (0.0527)
Population growth	-0.335* (0.192)	-0.443** (0.190)	-0.261 (0.210)	-0.176 (0.195)
TA _v trade openness	0.00101 (0.00962)	-0.00236 (0.00973)	-0.00315 (0.00921)	-0.00253 (0.00940)
TA _v dictatorship	-0.0976 (1.066)	0.873 (1.201)	0.523 (0.724)	0.247 (0.714)
TA _v education	-0.0317 (0.182)	-0.0957 (0.192)	-0.0865 (0.126)	-0.0211 (0.133)
TA _v health	-0.00544 (0.0249)	-0.0221 (0.0270)	0.0246 (0.0183)	0.0188 (0.0176)
Temperate zone	1.051 (1.092)	0.426 (0.966)	-0.446 (0.809)	-0.306 (0.821)
Observations	184	187	341	344
RMS Residual	2.841	2.796	3.707	3.652
R-squared
Number of countries	94	95	87	88

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.B.02(b): Investment rate, RE same samples with time averages of control variables

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	1.791** (0.886)	-0.00627 (0.817)	-0.381 (0.732)	-0.485 (0.683)
Rule of Law >Q3	-0.425 (0.798)	1.294 (0.838)	1.156 (0.941)	1.234 (0.872)
gy62	0.0211 (0.177)	-0.0218 (0.184)	0.000152 (0.186)	-0.0247 (0.183)
Investment rate	0.575*** (0.0691)	0.559*** (0.0722)	0.574*** (0.0709)	0.572*** (0.0714)
Population growth	-0.531** (0.236)	-0.656*** (0.239)	-0.572** (0.257)	-0.574** (0.248)
TA _v trade openness	-0.00377 (0.0101)	-0.00640 (0.0106)	-0.00608 (0.0106)	-0.00724 (0.0107)
TA _v dictatorship	0.210 (1.124)	1.260 (1.247)	1.059 (1.223)	1.192 (1.233)
TA _v education	-0.0720 (0.198)	-0.141 (0.209)	-0.149 (0.203)	-0.143 (0.210)
TA _v health	0.0139 (0.0265)	-0.00491 (0.0271)	-0.00760 (0.0257)	-0.00691 (0.0257)
Temperate zone	0.476 (1.211)	-0.221 (1.070)	-0.184 (1.161)	-0.196 (1.119)
Observations	163	163	163	163
R-squared
Number of countries	83	83	83	83
RMS Residual	2.625	2.573	2.613	2.590

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6.B.03(a): Investment rate, RE2 full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	0.726 (0.668)	0.0414 (0.774)	-0.408 (0.526)	0.0679 (0.498)
Rule of Law >Q3	-0.563 (0.492)	0.0529 (0.517)	-0.853* (0.487)	-0.492 (0.415)
gy62	-0.0125 (0.0939)	-0.0557 (0.0864)	0.00198 (0.0801)	-0.0297 (0.0859)
Investment rate	0.572*** (0.0541)	0.607*** (0.0632)	0.652*** (0.0444)	0.695*** (0.0422)
Population growth	-0.528*** (0.147)	-0.612*** (0.145)	-0.444*** (0.137)	-0.357*** (0.130)
Trade openness	-0.00825 (0.00558)	-0.00893 (0.00698)	-0.00305 (0.00572)	-0.00452 (0.00523)
Dictatorship	0.0605 (0.628)	0.619 (0.733)	-0.175 (0.501)	-0.456 (0.472)
Observations	277	294	442	428
RMS Residual	2.813	2.971	3.752	3.741
R-squared
Number of countries	152	161	130	119

Table 6.B.03(b): Investment rate, RE2 same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	0.905 (0.708)	0.380 (0.704)	-0.290 (0.637)	-0.537 (0.639)
Rule of Law >Q3	-0.182 (0.501)	0.330 (0.524)	0.222 (0.573)	-0.128 (0.509)
gy62	0.0887 (0.120)	0.0659 (0.122)	0.0568 (0.120)	0.0450 (0.122)
Investment rate	0.580*** (0.0624)	0.575*** (0.0624)	0.571*** (0.0619)	0.567*** (0.0627)
Population growth	-0.600*** (0.170)	-0.617*** (0.177)	-0.587*** (0.188)	-0.610*** (0.178)
Trade openness	-0.00874* (0.00516)	-0.0104** (0.00524)	-0.0109** (0.00534)	-0.0104* (0.00548)
Dictatorship	0.186 (0.691)	0.461 (0.733)	0.460 (0.738)	0.483 (0.746)
Observations	205	205	205	205
RMS Residual	2.629	2.616	2.616	2.604
Number of countries	108	108	108	108
R-squared

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

6. Rule of Law and economic growth: A panel data re-examination

Table 6.B.04(a): Investment rate, FE full samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	0.877 (0.977)	-0.154 (1.352)	-0.0506 (0.682)	-0.234 (0.680)
Rule of Law >Q3	-0.525 (0.722)	1.653* (0.898)	-0.0601 (0.673)	-0.938 (0.879)
gy62	0.0667 (0.113)	-0.0248 (0.111)	-0.00646 (0.0966)	-0.0760 (0.110)
Investment rate	0.227** (0.113)	0.221** (0.102)	0.183** (0.0785)	0.263** (0.118)
Population growth	-0.217 (0.297)	-0.459 (0.410)	0.00552 (0.370)	0.188 (0.314)
Trade openness	-0.0199 (0.0216)	-0.0112 (0.0211)	-0.0220 (0.0205)	-0.0315 (0.0195)
Dictatorship	0.677 (1.384)	0.662 (1.313)	-2.079* (1.194)	-1.700 (1.051)
Observations	277	294	442	428
Number of countries	152	161	130	119
RMS Residual	1.816	1.902	2.816	2.902
R-squared	0.088	0.091	0.058	0.088

Table 6.B.04(b): Investment rate, FE same samples

VARIABLES	(1) Heritage Foundation	(2) WGI	(3) ICRG	(4) Fraser Institute
Rule of Law <Q1	-0.887 (1.053)	-0.748 (0.639)	-1.124 (0.892)	-1.418** (0.636)
Rule of Law >Q3	-0.108 (0.954)	2.421 (1.462)	-0.756 (1.054)	0.729 (1.147)
gy62	0.0882 (0.121)	0.0876 (0.120)	0.0572 (0.114)	0.0632 (0.114)
Investment rate	0.0459 (0.110)	0.0353 (0.105)	0.0399 (0.113)	0.0421 (0.110)
Population growth	-0.854** (0.402)	-0.909** (0.381)	-0.943** (0.406)	-0.791** (0.369)
Trade openness	-0.0332 (0.0248)	-0.0253 (0.0232)	-0.0339 (0.0257)	-0.0288 (0.0253)
Dictatorship	-0.730 (1.319)	-0.910 (1.337)	-1.083 (1.423)	-0.718 (1.323)
Observations	205	205	205	205
R-squared	0.102	0.129	0.110	0.125
Number of countries	108	108	108	108
RMS Residual	1.519	1.496	1.513	1.500

Regressand leads regressors by one five-year period; fixed time effects included.

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

7 Conclusion

This dissertation provides a comprehensive analysis of the issues surrounding the construction and use of governance indicators. While chapter two gives an overview of the governance indicator market and background information on five of the most popular governance indicators, chapter three analyses the actual use and misuse of governance indicators. Chapter four analyses in-depth the most widely-used governance indicators, the Worldwide Governance Indicators from the World Bank Institute and chapter five attempts to explain their popularity and looks inside the organisation World Bank, one of the frontrunners in the supply of governance indicators, to gain a deep understanding of the politics surrounding the supply and use of governance indicators. In Chapter six, we investigate the relationship between Rule of Law and economic growth with several panel data methods and test the robustness of results depending on the choice of Rule of Law indicator.

7. 1 Sources of governance indicators (chapter two)

Following the boom in interest in the quality of governance over the last decade by international investors, aid donors and development analysts, the supply of governance indicators is rising. While at least 140 user-accessible sets of governance indicators are available nowadays, the use of governance indicators is concentrated around a few measurements. Most of these popular measurements have been available already before the “governance boom” or draw partly on

data whose origins precede the recent rise in interest in the quality of governance.

The non-governmental organisation Freedom House provides already for several decades their annual rating of political rights and civil liberties. The commercial International Country Risk Guide (ICRG) was created in the wake of the costly financial shock to international lenders caused by the fall of the Shah of Iran in 1979. Transparency International used the ICRG and other sources from commercial risk rating agencies for its famous “Corruption Perceptions Index” that helped to put governance on the agenda in the mid-nineties, followed by the more comprehensive set of indicators, the Worldwide Governance Indicators, which are also based on aggregating existing governance data. The International Development Association added over the last decade a strong governance component to its Country and Policy Institutional Assessments. World Bank staff constructs these ratings themselves and draw, among other sources of information, on the ratings mentioned above.

These popular governance indicators have two characteristics in common: first, they compile a large amount of information which they reduce to a single number for a given country in a given year. Second, they are perceptions-based indicators, i.e. they rely on the perceptions of households, business people, civil servants or experts of the quality of governance.

Next to the most popular indicators, there are many more indicators available and new initiatives continuously try to complement and improve upon existing measurements. Many users are not aware of this plethora of available data. This chapter takes stock of the major guides and databases that have emerged over the last years to help potential users of indicators find the right indicator and understand its strengths and weaknesses.

7.2 Uses of governance indicators (chapter three)

The primary direct users of governance indicators, besides journalists, are international investors, aid donors and academics.

International investors are interested because of the growth of their assets in developing and “emerging market” economies. They also want back-up to traditional macroeconomic-based country-risk indicators that have failed to predict costly financial crises. Banks and Multinationals as well as traditional risk rating agencies increasingly include governance indicators in their country risk ratings, which influence investment decisions in developing countries. The importance of the country risk ratings is rising since Basel II: While risk weights, which determine the amount a bank has to hold in reserve for a loan as a back-up, were assigned to countries according to the simple classification OECD or non-OECD country with Basel I, risk weights depend since Basel II on country risk ratings.

Bi- and multilateral donors increasingly rely on governance indicators for allocating aid to developing countries. To illustrate, The US Millennium Challenge Corporation allocates billion of dollars on the basis of governance indicators such as the Worldwide Governance Indicators and Freedom House indicators and the governance indicators of the Country and Policy Institutional Assessment play a crucial role for the World Bank's International Development Association's lending decisions.

Academics use governance indicators frequently in their quantitative studies. Some of these studies provide evidence for a positive impact of good governance on economic growth (e.g. Acemoglu, Johnson and Robinson, 2001, Kaufmann and Kraay, 2002b, Knack and Keefer, 1995) and on aid effectiveness (e.g. Burnside and Dollar, 1997, 2004). They have been crucial in convincing bi- and multilateral organisations to shift the focus of development policies to the quality of governance.

Unfortunately, these three major groups of users often fail to take into account the limitations of the indicators they use, such as lack of comparability over time and across countries and conceptual weaknesses. Furthermore, besides local media, the active users of governance indicators for decision-making purposes are overwhelmingly external stakeholders, i.e. foreign investors, development analysts and donors. 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.

Two implications for policy-makers and academia emerge. First, a better understanding of the limitations of existing data is necessary to use indicators correctly (see chapter four). Given the rise in importance of indicators for far-reaching decisions about resource allocations to developing countries, a more careful choice and use of indicators is of utter importance. Second, the current challenge is to construct 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. In a few fields, such indicators have been already developed - and their number is rising (see chapter five)

7. 3 The Worldwide Governance Indicators (chapter four)

The last chapter demonstrated the rising importance of governance indicators for analysis and decision-making purposes. Given the high stakes for developing countries, this chapter provides an in-depth analysis of the most popular indicators to find out whether decisions are made on the basis of sound data.

The Worldwide Governance Indicators (WGI) are widely used to compare the quality of governance over time and across countries, for aid-allocation decisions, for risk ratings, for academic analysis and for media articles. Constructed by the World Bank Institute, they summarize data from more than 30 expert assessments and household

and firm surveys, and are available from 1996 to today for more than 200 countries.

Unfortunately, users of these indicators frequently ignore five interrelated problems with these indicators. First, disturbances are likely to be correlated which puts the consistency and unbiasedness of the estimates into question and increases the size of the standard errors. Second, WGIs lack transparency because of their high complexity, lack of reproducibility and lack of underlying theory of governance. Third, the WGIs are based on a weak conceptual foundation. Fourth, the WGIs suffer from sample bias which is difficult to decipher for users. To illustrate, the aggregation procedure assigns 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. Many users are not aware of this difference in weight and its consequences. Fifth, the WGIs cannot be used for direct comparisons of the quality of governance over time.

The ignorance of the limitations of the WGIs has severe implications for policy-making and academic research. First, resource allocation decisions that ignore the WGI's lack of comparability across countries and over time may be inefficient – and unfair. Second, the indicators may give incentives for wrong governance reform to developing countries. Third, governments are blamed in some cases for results they are not responsible for.

Accordingly, the WGI is useful as a first snapshot of a country's quality of governance, but should not be used in any automatic way for decision-making purposes. The underlying sources provide users with a wealth of information about a country's quality of governance. Contradictions in their assessments of a country's quality of governance should not be ignored. Instead, they help to gain a deeper understanding of differences in various aspects of governance and the subjectivity of assessments.

7.4 Politics of governance ratings (chapter five)

The World Bank plays a leading role in the supply of cross-country comparable governance indicators, of both aggregated summary measurements (e.g., the WGI), firm surveys (e.g., BEEPs) and more recent, "actionable" indicators (e.g., PEFA). The World Bank produces governance ratings for several interrelated reasons. First, ratings could help identify areas of reform and measure the success of governance reforms. Second, governance indicators help to put governance on the agenda and create incentives for developing countries to improve their governance. Third, governance indicators can enhance country dialog, if correctly used. Fourth, indicators can contribute to a higher transparency of aid-allocation decisions. Fifth, indicators can be used for quantitative analysis. Sixth, a number of institutional reasons drive the production of governance indicators.

Due to the limitations of the first generations of indicators to fulfil these high expectations, the World Bank currently focuses on constructing so-called “actionable” indicators, i.e., indicators that are informative about what reform countries could undertake to improve their ratings and that track progress over time. However, progress is slow because it is not clear what is “action worthy”. The discussion about “actionable” indicators and about the limitations of “best-practice” indicators currently triggers a deep debate about what governance is good for developing countries at different stages of development. This current controversy about the use of existing indicators and future developments will heavily impact World Bank policies in developing countries, not the least because indicators determine aid-allocation and influence governance reforms.

Outside the World Bank, the most popular World Bank governance indicators are the WGIs. While within the World Bank staff is very well aware of their limitations, outside users rely on them for far reaching decisions and judgment about country’s quality of governance, despite the availability of more meaningful and robust indicators. A combination of factors goes far to explain the domination of WGIs outside the World Bank: there is a high demand for summary measurements of the quality of governance with a broad country coverage and users often lack knowledge about the limitations of the indicators. They also perceive a lack of alternatives and show signs of herd behaviour.

Several recommendations for policy-makers and development researchers emerge. In the short-run, a plethora of data is publicly accessible and explained in governance indicator guides and inventories. Furthermore, producers can improve the way they advertise their indicators and enhance their transparency to minimize misunderstandings about the meaning and robustness of the indicators.

In the medium to long run, more specific, meaningful, and “actionable” indicators will be developed. In order to guarantee that these indicators actually benefit developing countries, transparency and “actionability” are not sufficient. Most crucial is better knowledge about what is “action worthy”, that is, what governance reforms would reduce poverty and trigger sustained economic growth. Currently a number of development researchers and practitioners express their concern that the “good governance agenda” is overloaded and that different institutions than those promoted in the most popular indicators might be necessary for inclusive growth. The questions and conflicts arising from the process of developing transparent “actionable” indicators can serve as a basis for debate, together with more qualitative research, about what governance is conducive to development, depending on the country context. It is much better to have this debate now than to start (as happened with so many other development fashions) implementing governance reforms on a large scale that, despite good intentions, might hinder country development more than they help it. More robust and transparent indicators could contribute to research on the relationship between governance and growth in developing countries, because they would allow for

identifying sequencing issues in specific reforms and relationships between improvements in specific governance areas⁷⁷.

7.5 Rule of Law and economic growth: A panel data re-examination (chapter six)

This chapter is both a practical application of the use of Governance Indicators and a partial test of the widespread belief that good governance, as commonly measured, triggers economic growth. The latter claim often serves as a justification for the proliferation of Governance Indicators.

We test the robustness of the effect of one of the core aspects of governance, the Rule of Law, on economic growth with different panel data methods. We rely on a comprehensive panel data set which includes several Rule of Law variables that are available for up to four decades: the Heritage Foundation indicator “Property Rights”, the Worldwide Governance Indicator “Rule of Law”, the Fraser Institute Indicator “Legal Structure and Security of Property Rights” and the International Country Risk Guide Indicator “Law and Order”. All four indicators focus mainly on the private sector.

Any findings of a positive effect of the Rule of Law on economic growth are not robust. Results vary heavily with respect to

⁷⁷ For a more detailed discussion of alternatives and future trends, see also Arndt and Romero (2008) and Arndt and Oman (2006)

the choice of Rule of Law indicators, the panel data method, the sample period and the scaling of the Rule of Law indicator. The Heritage Foundation indicator never has a significant effect in the expected direction. The ICRG Indicator only shows marginally significant effects when categorical dummies are introduced and practically no significant effects when the Rule of Law indicators are left in their original cardinal scale. The WGI and Fraser more frequently show significant effects. Differences in the effect of Rule of Law proxies are in some cases due to the choice of sample: Every Rule of Law indicator is available for different countries and years. However, differences between the WGIs and the Fraser on the one hand and the Heritage Foundation and the ICRG indicator on the other hand mostly remain when a common sample is chosen.

In those regressions that allow for full country heterogeneity, i.e. in fixed effects regressions, the Rule of Law proxies are barely significant. Some significant effects of the ICRG and the Fraser indicator are not robust.

The significance of some of the control variables such as trade openness varies with the choice of the Rule of Law indicator, which is partly but not completely due to differences in the sample.

Findings have the following implication for policy-makers: Claims that improving the Rule of Law, as defined by popular indicators, should be the policy priority if the policy objective is to foster economic growth, cannot be sustained at present by

econometric panel data analysis. A lack of evidence for a systematic effect of the Rule of Law on economic growth does of course not mean that governance and the Rule of Law do not matter to trigger and maintain economic growth. The Rule of Law is a complex and not uniquely defined concept, as is its relationship with economic growth. Better indicators and more refined in-depth country studies are necessary to gain a deeper understanding of the type of Rule of Law systems and reforms that are conducive to economic growth. Policy advice would need to take into account the level of development and the local context.

7.6 The governance indicator paradox

This dissertation may leave the reader with the following, somehow paradoxical thought: Why should there be apparently rising demand for Governance Indicators that appear to be of so little policy value? In fact, two key points may help to clarify the conclusion drawn from this dissertation and to answer this question. First, many existing Indicators as well as indicators currently under development are valuable for policy makers as long as, second, their meaning is not overstretched and they do not replace the decision making process.

High quality indicators are transparent, clearly defined and meaningful. Their methodology and assumption are clearly exposed and all the underlying data is publicly available unless there are extremely good reasons not to disclose the data. Producers of high quality indicators should advertise them carefully and clearly expose

their limitations and the purposes they can be used for. These characteristics and procedures allow potential users of indicators to fully understand the strengths and the weaknesses of the indicators.⁷⁸

High quality indicators can be useful for clarifying the underlying issues, as quantification forces to be specific. Useful time series data can help identifying sequencing issues in governance reforms across countries and show changes over time. Relevant indicators can also help to group countries according to their governance systems.

Obviously, no indicator can ever be perfect. Choices of different methods (e.g. perception vs. facts-based) entail trade-offs (e.g. measuring quality of actual practices vs. higher “objectivity”) that cannot be avoided. Indicators can never show the full picture, they are by definition a simplification of a complex reality.

The real problem arises if indicators are asked to deliver more than they can. Policy makers in need for justification of their action may overstretch the meaning of the indicators, draw simplistic conclusions from them and rely on them for automatic decision making. This is a way for policy makers to avoid taking responsibility for their decisions and the uncertainty associated with any decision by referring to apparently “objective” criteria. As Van de Walle and

⁷⁸ See “Handbook on Constructing Composite Indicators: Methodology and User Guide” (OECD and EC 2008) for guidance on the construction of high quality composite indicators.

Roberts phrase it, “the task of imposing order on a complex reality is often difficult” (Van de Walle, Roberts, p. 223). “Indicators have created for policy makers and citizens an appearance or illusion of control” (Van de Walle, Roberts, 2008, p. 222).

There is a paradox that in an increasingly complex world increasingly simplistic tools are used for decision making. Transparent indicators can contribute to provide information in a systematic way. They cannot and should not replace a political decision making process that takes into account the specific context.

7.7 Limitations and recommendations for future work

The current study has several limitations that lead to recommendations for future work around governance indicators. First, chapter two to five draw on a limited survey of staff of bi- and multilateral aid organisations, the private sector and academics due to time- and financial constraints. Second, more systematic information on the views of developing countries’ governments and the impact of different kinds of governance ratings’ on developing countries policy-making and the quality of governance would have been useful. Finally, the panel data results of chapter six are of limited use for reform advice to policy-makers, because every governance indicator used in a regression is highly correlated with other governance indicators. No instrument for Rule of Law indicators is available that is clearly uncorrelated with other governance indicators in the error term and it is not possible to include all governance variables due to

multicollinearity issues. It is therefore difficult to use any results of a positive or negative impact of the quality of the Rule of Law on growth for reform advice.

A number of research projects would be beneficial to overcome these limitations. First, more research on the actual use of indicators, in particular by “internal” stakeholders, would be helpful to evaluate the usefulness of indicators and help improving them. Second, an analysis of the actual impact of the use of governance indicators on the quality of governance in developing countries would contribute to the same aim⁷⁹. Third, comprehensive studies on the use, impact, and usefulness of “actionable” indicators are not available yet, because such indicators are a relatively new phenomenon. Such studies would be crucial to inform new developments in the field. Fourth, in-depth country studies and a better methodology to assess the quality of governance in developing countries is necessary to move away from standard best-practice advice to a more tailored advice that takes into account the local context and growth constraints⁸⁰.

⁷⁹ Currently, the Graduate School of Governance undertakes such research on Mozambique. A project with the University of Oxford is planned to develop a methodology that allows for comparing the impact of international ratings across countries.

⁸⁰ The Maastricht Graduate School of Governance is collaborating with the French Development Agency (AFD) and researchers from universities around the world to develop such a methodology.

References

- Aberbach, J. D. & Rockman, B. A. (2002). Conducting and Coding Elite Interviews. *Political Science & Politics*, 35(4), 673-676.
- Acemoglu, D., Johnson, S. & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 1(5), 1369-1401.
- Alesina, A. & Rodrik, D. (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109, pp. 465-490.
- Albouy, D. (2006). *The Colonial Origins of Comparative Development: An Investigation of the Settler Mortality Data*. Unpublished. University of California at Berkeley.
- Alonso, W. & Starr, P. (1987). *The Politics of Numbers*. New York: Russell Sage Foundation.
- Anderson, J. (2007). *Comments on the first draft note of Arndt and Romero (Arndt & Romero 2008)*. Unpublished.
- Andrews, M. (2008). *Are One-Best-Way Models of Effective Government Suitable for Developing Countries?* Faculty Research Working Papers Series. Cambridge, MA: John F. Kennedy School of Government - Harvard University. Retrieved May 2008, from <http://ksgnotes1.harvard.edu/Research/wpaper.nsf/rwp/RWP08-014>
- Arndt, C. (2008). The Politics of Governance Ratings. *International Public Management Journal*, Volume 11, Issue 3, pp. 275-297.
- Arndt, C. & de Crombrugghe, D. (2008). *Rule of Law and Economic Growth: A panel data re-examination*. Unpublished Manuscript.
- Arndt, C. & Oman, C. (2006). *Uses and Abuses of Governance Indicators*. Development Centre Studies. Paris: OECD.
- Arndt, C., Knack, S. & Oman, C. (2007). *A reply to Kaufmann, Kraay and Mastruzzi's „Answering the critics“*. Unpublished manuscript.
- Arndt, C. & Romero, C. (2008). *Review of the policy utility of the Worldwide Governance Indicators 2007 for the Central American Countries*. World Bank Working Paper Series on Public Sector Management. Washington, D.C.: World Bank.

- Aron, J. (2000). Growth and Institutions. A Review of the Evidence. *The World Bank Research Observer* 15(1), 99-135.
- Arrunada, B. (2007). *Pitfalls to Avoid when Measuring Institutions: Is Doing Business Damaging Business?* Barcelona: Pompeu Fabra University.
- Barro, R. (1991). Economic Growth in a Cross-Section of Countries. *Quarterly Journal of Economics* 106(2), 407-443.
- Barro, R. & Lee, J.W (2000). International data on educational attainment: Updates and Implications. *Oxford Economic Papers* 3, 541-563.
- Barro R. & Sala-i-Martin X. (2004). *Economic Growth (2nd ed.)*. Cambridge: MIT Press.
- Benhabib, J. & Spiegel, M.M. (1994). The Role of Human Capital in Economic Development: Evidence from aggregate cross-country data. *Journal of Monetary Economics* 34(2), 143-173.
- Berg, J. & Cazes, S. (2007). *The Doing Business Indicators: Measurement issues and political implications*. Economic and Labour Market Paper 2007/6. Geneva: International Labour Organisation. May 2008, from <http://www.ilo.org/public/english/employment/download/elm/elm07-6.pdf>
- Berry, J. M. (2002). Validity and Reliability Issues in Elite Interviewing. *Political Science and Politics*, 35(4), 679-682.
- Berthélemy, J.-C. & Tichit, A. (2004). Bilateral Donors' Aid Allocation Decisions — A Three-Dimensional Panel Analysis. *International Review of Economics & Finance, Elsevier*, 13(2004), 253-274.
- Besançon, M. (2003). Good Governance Rankings: The Art of Measurement. *World Peace Foundation Reports*, 36.
- Bevan, G. & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public Administration* 84(3), 517-538.
- Bouchet, M. H., Clarke, E. & Gros Lambert, B. (2003). *Country Risk Assessment: A Guide to Global Investment Strategy*. Wiley Finance Series. West Sussex: John Wiley & Sons.

- Bouckaert, G. & Halligan, J. (2008). *Managing performance: international comparisons*. London and New York: Routledge/Taylor and Francis.
- Boyne, G. A., Meier, K. J., O'Toole, L. J. & Walker, R. M. (2006). *Public service performance: perspectives on measurement and management*. Cambridge: Cambridge University Press.
- Brock, W. & Durlauf S. (2001). Growth Empirics and Reality. *World Bank Economic Review* 15, 2, 229-272.
- Burnside, C. & Dollar, D. (1997). Aid, Policies, and Growth. *World Bank Policy Research Working Paper*, No. 1777. Washington, D.C.: The World Bank.
- Burnside, C. & Dollar, D. (2004). *Aid, Policies, and Growth: Revisiting the Evidence*. World Bank Policy Research Working Paper, No. 3251. Washington, D.C.: World Bank. Retrieved February 2008, from http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2004/04/21/000009486_20040421103444/Rendered/PDF/wps3251Aid.pdf
- Business and Industry Advisory Committee to the OECD. Investment - BIAC Position on Incentive. Global Forum on International Investment, 5-6 December 2002, Shanghai.
- Caisse des Depots et Consignations (1998). *Emerging Regions Report*, 5(4), 70-71.
- Carothers, T. (2003). Promoting the Rule of Law Abroad. The Problem of Knowledge. *Rule of Law Series. Democracy and Rule of Law Project*. Number 34. Retrieved May 2008, from <http://www.carnegieendowment.org/publications/index.cfm?fa=view&id=1169>
- Charles, C., Dijakovic, V., Kaempfer, M. & Kozlova, M. (2008). *Indicateurs de gouvernance. Une étude comparative des approches et des stratégies de la Banque mondiale, de l'Union européenne et du PNUD*. Rapport final - Atelier international 2007/2008. Paris: Institut d'Etudes Politiques de Paris.

- Cheibub, J. A. & Gandhi, J. (2004). *Classifying Political Regimes: A Sixfold Classification of Democracies and Dictatorships*. Paper presented at the Annual Meeting of the American Political Science Association.
- Chong, A. & Calderon, C. (2000). Causality and Feedback Between Institutional Measures and Economic Growth. *Economics and Politics*, 12(1), 69-81.
- Cornia, G. A. & Popov, V. (Eds.). (2001). *Transition and Institutions: The Experience of Gradual and Late Reformers*. Oxford: Oxford University Press.
- Davis, K. E. (2004). *What Can the Rule of Law Variable Tell Us About Rule of Law Reforms?* NYU Law and Economics Research Paper No. 04-026
- Dexter, L.A. (2006). *Elite and Specialized Interviewing. With a new introduction by Alan Ware and Martín Sánchez-Jankowski*. University of Essex, Colchester/UK: ECPR Press.
- DiMaggio, P. J. & Powell, W. W. (1991). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organisational Fields. In W. Powell & P. DiMaggio (Eds.), *The New Institutionalism in Organizational Analysis* (pp. 63-82). London and Chicago: The University of Chicago Press.
- Djankov, S., Darshini, M., McLiesh, C. & Ramalho, R. (2005). *Doing Business Indicators: Why Aggregate, and How to Do It*. Washington D.C.: World Bank. Retrieved June 2008, from http://siteresources.worldbank.org/EXTAFRSUMAFTPS/Resources/db_indicators.pdf
- Djankov, S. (2008). *A Response to Is Doing Business Damaging Business*. Washington D.C.: World Bank.
- Dollar, D. & Levin, V. (2004). *The Increasing Selectivity of Foreign Aid (1984-2002)*. World Bank Policy Research Working Paper, No. 3299. Washington, D.C.: World Bank.
- Dunleavy, P. (1991). *Democracy, bureaucracy and public choice: economic explanations in political science*. New York: Harvester Wheatsheaf.

- Easterly, W. & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112(4, November):1203–50.
- Easterly, W. (2002). *The Elusive Quest for Growth*. Cambridge, Massachusetts: MIT Press.
- Ellickson, R. C. (2001). The market for social norms. *American Law and Economics Review*, 3(1), 1-49.
- Eviatar, D. (2003, July 23). Do Aid Studies Govern Policies or Reflect Them? *New York Times*.
- Frankel J. & Romer, D. (1999). Does Trade Cause Growth? *American Economic Review* 89(3), 379-399.
- Gallup J.L., Mellinger A. D. & Sachs J. D. (1999). *Geography Datasets*. Harvard Centre for International Development. Retrieved December 2007 from: <http://www.cid.harvard.edu/ciddata/geographydata.htm>
- Galtung, F. & Sampford, C. (2005). Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices. In F. Galtung & C. Sampford (Eds.), *Measuring Corruption*. London: Ashgate
- Gelb, A., Ngo, B. & Ye, X. (2004). *Implementing Performance-Based Aid in Africa: The Country Policy and Institutional Assessment*. World Bank Africa Region Working Paper Series, No. 77. Washington, D.C.: World Bank.
- Girvan, N. (2002). *Problems with UNDP Governance Indicators*. Retrieved June 2007, from <http://www.southcentre.org/info/southbulletin/bulletin41/bulletin41-04.htm>
- Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F. & Shleifer, A. (2004). Do Institutions Cause Growth? *Journal of Economic Growth*, 9(3), 271-303.
- Global Integrity (2007). *Assessments and Data*. Retrieved June 2007, from <http://www.globalintegrity.org/data/index.cfm>
- Globerman, S. & Shapiro, D. (2002). Global Foreign Direct Investment Flows: The Role of Governance Infrastructure. *World Development*, 30(11), 1899-1919

- Globerman, S. & Shapiro, D. (2003). Governance Infrastructure and U.S. Foreign Direct Investment. *Journal of International Business Studies*, 34(14), 61-74.
- Goldstein, K. (2002). Getting in the Door: Sampling and Completing Elite Interviews. *Political Science & Politics*, 35(4), 669-672.
- Gray, H. (2007). *Governance for Economic Growth and Poverty Reduction: Empirical Evidence and New Directions Reviewed*. Retrieved July 2008, from <http://www.gsdr.org/docs/open/RET422.pdf>
- Griffith-Jones, S. & Gottschalk, R. (2006). Cost of Currency Crises and Benefits of International Financial Reform. In I. Kaul (Ed.), *The New Public Finance: Responding to Global Challenges*. New York, NY: Oxford University Press.
- Grindle, M. S. (2007). Good Enough Governance Revisited. *Development Policy Review* 25(5), 553-574.
- Guha, K. (2007a, July 10). "Countries unhappy at World Bank ratings". *Financial Times.com*. Retrieved August 2008, from <http://www.ft.com/cms/s/0/46117ad8-2f07-11dc-b9b7-0000779fd2ac.html>

- Guha, K. (2007b, July 12). "World Bank ratings spark anger". *Financial Times.com*. Retrieved August 2008, from <http://www.ft.com/cms/s/0/34a681d2-2f4b-11dc-b9b7-0000779fd2ac.html>
- Gwartney, J. & Lawson, R. (2006). *Economic Freedom of the World: 2006 Annual Report*. Vancouver: The Fraser Institute.
- Gwartney, J. & Lawson, R. with Norton, S. (2008). *Economic Freedom of the World: 2008 Annual Report*. The Fraser Institute.
- Heritage Foundation (1995). *Index of Economic Freedom 1996*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (1996). *Index of Economic Freedom 1997*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (1997). *Index of Economic Freedom 1998*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (1998). *Index of Economic Freedom 1999*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (1999). *Index of Economic Freedom 2000*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (2000). *Index of Economic Freedom 2001*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (2001). *Index of Economic Freedom 2002*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (2002). *Index of Economic Freedom 2003*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (2003). *Index of Economic Freedom 2004*. Washington, D.C.: The Heritage Foundation and Dow Jones &

- Company, Inc.
- Heritage Foundation (2004). *Index of Economic Freedom 2005*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (2005). *Index of Economic Freedom 2006*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Heritage Foundation (2006). *Index of Economic Freedom 2007*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Herman, B. (2004). *How Well Do Measurements of An Enabling Domestic Environment For Development Stand Up*. Geneva: UNCTAD. Retrieved February 2008, from www.g24.org/003gva04.pdf
- Hermes, N. & Lensink, R. (2001). Changing the Conditions for Development Aid: A New Paradigm? *Journal of Development Studies* 37(6), 1-16.
- Heston, A., Summers, R. & Aten, B. (2006). *Penn World Table Version 6.2*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Holmes, K. R., Feulner, E. J., O'Grady, M. A. & et al. (2008). *Index of Economic Freedom 2008*. Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc.
- Hood, C. (2007). Public service management by numbers: why does it vary? Where has it come from? What are the gaps and the puzzles? *Public Money and Management*, 27(2), 95-102.
- Hood, C., Beeston, C. & Dixon, R. (2008). Rating the Rankings: Assessing International Rankings of Public Service. *International Public Management Journal*.
- Huther, J. & Shah, A. (1996). *A Simple Measure of Good Governance*. Washington, D.C.: World Bank.
- Huther, J. & Shah, A. (1998). *Applying a Simple Measure of Good Governance to the Debate on Fiscal Decentralisation*. Policy Research Working Paper No. 1894. Operations Evaluation

- Department. Country Evaluation and Regional Relations Division. Washington, D.C.: World Bank.
- Institute of International Finance (2005). *Capital Flows to Emerging Market Economies*. Washington D.C.: World Bank.
- Inter-American Development Bank (2007). *DataGob: Governance Indicator Database*. Retrieved June 2007, from <http://www.iadb.org/datagob/>
- International Finance Corporation (1998). *Emerging Stock Markets Factbook*. Washington, D.C.: World Bank.
- International Monetary Fund (2006). *Albania: Poverty Reduction Strategy Paper—Annual Progress Report*. Retrieved June 2007, from <http://www.imf.org/external/pubs/ft/scr/2006/cr0623.pdf>
- International Monetary Fund (2006). *Country Report on Albania* No. 06/285. Washington D.C.: International Monetary Fund.
- Iqbal, K. & Shah, A. (2006). *Truth in advertisement: How do Governance Indicators Stack Up?* Unpublished Paper. Washington, DC: World Bank.
- Iqbal, K. & Shah, A. (2008). *How Do Worldwide Governance Indicators Measure Up?* Washington, D.C.: World Bank. Retrieved May 2008, from <http://siteresources.worldbank.org/PSGLP/Resources/Howdoworldwidegovernanceindicatorsmeasureup.pdf>
- Jabes, J. (2002). *On the (F)utility of Governance Indicators: Lessons from Countries in Transition*. Speech given as part of "Assessing Good Governance" panel at the International Institute of Administrative Sciences Second Specialized International Conference. November 2002.
- Johnson, D. & Zajonc, T. (2006). *Can Foreign Aid Create an Incentive for Good Governance?* Evidence from the Millennium Challenge Corporation. Working Paper, Harvard University.
- Jüttner, J. D. & McCarthy, J. (2000). *Modelling a Rating Crisis*. Macquarie University. Unpublished.
- Kalaitzidakis, P. , Mamuneas, T. & Stengos, T. (2000). A Non-Linear

- Sensitivity Analysis of Cross-Country Growth Regressions. *Canadian Journal of Economics* 33, 3, 604-617.
- Kaufmann, D. (2004). Corruption, Governance and Security: Challenges for the Rich Countries and the World. In *Global Competitiveness Report 2004/2005*. New York, NY: Palgrave Macmillan Ltd.
- Kaufmann, D. (2003). *Rethinking Governance, Empirical Lessons Challenge Orthodoxy*. Macroeconomics 0308007, Economics Working Paper Archive EconWPA, World Bank, Washington, D.C. Retrieved February 2008, from www.worldbank.org/wbi/governance/pdf/rethink_gov_stanford.pdf#search='%E2%80%9CRethinking%20
- Kaufmann, D. & Kraay, A. (2002a). *Governance Indicators, Aid Allocation and the Millennium Challenge Account*. Washington, D.C.: World Bank. Retrieved February 2008, from www.worldbank.org/wbi/governance/pdf/mca_summary.pdf
- Kaufmann, D. & Kraay, A. (2002b). Growth Without Governance. *Economía*, 3(1), 169-215.
- Kaufmann, D. & Kraay, A. (2003). *Governance Matters III: New Indicators for 1996-2002 and Addressing Methodological Challenges*. World Bank Policy Research Working Paper, No. 3106. Washington, D.C.: World Bank. Retrieved February 2008, from www.worldbank.org/wbi/governance/pdf/govmatters3_synth.pdf
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2003). *Governance Matters III: governance indicators for 1996-2002*. Washington: World Bank.
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2005a). *Illustrative Composite Governance Indicator Worldwide Map*. Washington, D.C.: World Bank. Retrieved February 2006, from <http://info.worldbank.org/governance/kkz2004/maps.html>
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2005b). *Governance Matters IV: Governance Indicators for 1996-2004*. World

Bank Policy Research Working Paper, No. 3630. Washington D.C.: World Bank Institute.

- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2006). *Governance Matters V: Governance Indicators for 1996-2005*. Retrieved February 2008, from http://siteresources.worldbank.org/INTWBIGOVANTCOR/Resources/1740479-1150402582357/2661829-1158008871017/gov_matters_5_no_annex.pdf
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2007a). *Governance Matters VI: Aggregate and Individual Governance Indicators for 1996-2006*. Retrieved February 2008, from <http://info.worldbank.org/etools/MDK21385805pagePK64257043piPK/pdf/govmatters6.pdf>
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2007b). *The Worldwide Governance Indicators Project: Answering the Critics*. World Bank Policy Research Working Paper 4149. Retrieved May 2008, from http://siteresources.worldbank.org/INTWBIGOVANTCOR/Resources/1740479-1149112210081/2604389-1167941884942/Answering_Critics.pdf
- Kaufmann, D., Kraay, A. & Zoido-Lobaton, P. (1999a). *Governance Matters*. World Bank Policy Research Working Paper, No. 2196. Washington, D.C.: World Bank Institute. Retrieved February 2008, from www.worldbank.org/wbi/governance/pubs/govmatters.html
- Kaufmann, D., Kraay, A. & Zoido-Lobaton, P. (1999b). *Aggregating Governance Indicators*. World Bank Policy Research Working Paper, No. 2195. Washington, D.C.: World Bank. Retrieved February 2008, from www.worldbank.org/wbi/governance/pubs/agginicators.html
- Kaufmann, D., Kraay, A. & Zoido-Lobaton, P. (2000). *Governance Matters: From Measurement to Action*. *Finance & Development*, 37(2), 10-13.
- Kaufmann, D., Kraay, A. & Zoido-Lobaton, P. (2002). *Governance Matters II: Updated Indicators for 2000/01*. World Bank Policy Research Working Paper, No. 2771. Washington, D.C.: World Bank. Retrieved February 2008, from <http://www->

- wds.worldbank.org/servlet/WDSCContentServer/WDSP/IB/2002/02/27/000094946_02020704044576/Rendered/PDF/multi0page.pdf
- Kaufmann, D. & Vicente, P. C. (2005). *Legal Corruption*. University of Oxford, Centre of Study of African Economies. Retrieved February 2008, from <http://ssrn.com/abstract=829844>
- King, R. & Levine, R., (1993). Finance and growth: Schumpeter might be right, *Quarterly Journal of Economics* 108 (3), 717-737.
- Knack, S. (2002). *Governance and Growth: Measurement and Evidence*. Forum Series on the Role of Institutions in Promoting Economic Growth. Washington, D.C.: USAID.
- Knack, S. (2006). *Governance and Growth*. Retrieved May 2008, from http://www.sdnpsd.org/sdi/issues/governance/governance/summarynoteg_governance_and_growth.pdf
- Knack, S. (2007). Measuring Corruption: A Critique of Indicators in Eastern Europe and Central Asia. *Journal of Public Policy*, 27(3), 255-291.
- Knack, S. & Keefer, P. (1995). Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures. *Economics and Politics* 7, 207-227.
- Knack, S. & Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *Quarterly Journal of Economics* 112 (4), 1251-1288.
- Knack, S., Kugler, M. & Manning, N. (2003). Second-Generation Governance Indicators. *International Review of Administrative Sciences*, Vol. 69, 345-364.
- Kraay, A. & Nehru, V. (2004). *When is External Debt Sustainable?* World Bank Policy Research Working Paper, No. 3200. Washington, D.C.: World Bank Institute. Retrieved February 2008, from http://www-wds.worldbank.org/servlet/WDSCContentServer/WDSP/IB/2004/06/02/000009486_20040602115347/Rendered/PDF/wps3200externaldbt.pdf

- Kurtz, M. & Schrank, A. (2007). Growth and Governance: Models, Measures, and Mechanisms. *The Journal of Politics*, 69(2), 538-554.
- Landman, T. & Häusermann, J. (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). Colchester: University of Essex, Human Rights Centre.
- Lee, K., Pesaran, M. & Smith, R., (1997). Growth and Convergence in Multi-Country Empirical Stochastic Solow Model. *Journal of Applied Econometrics* 12 (4) 357-392.
- Levine, R. & Renelt, D. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *The American Economic Review*, 82(4), 942-963.
- Mankiw N.G., Romer, D. & Weil, D., (1992). A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 107 (2), 407-437.
- Mauro, P. (1995). Corruption and Growth. *Quarterly Journal of Economics*, 110(3), 681-712.
- Meisel, N. (2004). *Governance Culture and Development: A Different Perspective on Corporate Governance*. Development Centre Studies. Paris: OECD.
- Metagora. (2007). *Inventory of Initiatives aimed at measuring human rights and democratic governance*. Retrieved June 2007, from http://www.metagora.org/html/aboutus/about_inventory.html
- Munck, G. L. & Verkuilen, J. (2002). Conceptualizing and Measuring Democracy: Evaluating Alternative Indices. *Comparative Political Studies*, Vol. 35(1), 5-34.
- Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1), 69-85.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- North, D. C. (2005). *Understanding the Process of Economic Change*. New Jersey: Princeton University Press.

- OECD / European Commission, Joint Research Centre (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris : OECD Publishing.
- Oman, C. P. & Wignaraja, G. (1991). *The Postwar Evolution of Development Thinking*. London: Palgrave MacMillan Ltd. and Development Centre Study.
- Oman, C. P. (2000). *Policy Competition for Foreign Direct Investment: A Study of Competition Among Governments to Attract FDI*. Paris: Development Centre Studies, OECD
- Ortiz, G. (2003). Latin America and the Washington Consensus: Overcoming Reform Fatigue. *Finance & Development, Sept. 2003*.
- Panizza, U. G. & Lora, E. A. (2002). Structural Reforms in Latin America Under Scrutiny. Text for Seminar “Reforming Reforms”, 11 March 2002, Fortaleza, Brazil. Retrieved February 2008, from <http://ssrn.com/abstract=311599>
- PEFA Secretariat (2008). PFM Assessment Reports. Retrieved August 2008, from <http://www.pefa.org>
- Political Risk Services (2007). *International Country Risk Guide (ICRG)*. Retrieved December 2007 from: <http://www.prsgroup.com/>
- Pollitt, C. (forthcoming). Simply the best? The international benchmarking of reform and good governance. In J. Pierre & P. Ingraham (Eds.), *Public sector administrative reform and the challenges of effective change*. Montreal: McGill-Queens University Press.
- Porter, T. M. (1995). *Trust in Numbers: the Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.
- Przeworski, A., Alvarez, M. E., Cheibub, J. A. & Fernando, L. (2000). *Democracy and Development: Political Institutions and Material Well-Being in the World, 1950-1990*. New York: Cambridge University Press.
- Rao, H., Monin, P. & Durand, R. (2003). Institutional change in Toque Ville: nouvelle cuisine as an identity movement in French gastronomy. *American Journal of Sociology*, 108(4),

795-843.

- Rao, H., Durand, R. & Monin, P. (2005). Border crossing: bricolage and the erosion of categorical boundaries in French gastronomy. *American Sociological Review*, 70(6), 968-991.
- Razafindrakoto, M. & Roubaud, F. (2006). *Are international databases on corruption reliable? A comparison of expert opinion surveys and household surveys in sub-Saharan Africa*. Retrieved February 2008, from http://www.dial.prd.fr/dial_publications/PDF/Doc_travail/2006-17_english.pdf
- Reisen, H. (2003). *Ratings Since the Asian Crisis*. Development Centre Working Paper No. 214, Paris: OECD. Retrieved July 2008, from www.oecd.org/dataoecd/22/8/1934625.pdf
- Rigobon R. & Rodrik D. (2004) *Rule of Law, Democracy, Openness, and Income: Estimating the Interrelationships*, NBER Working Papers 10750. National Bureau of Economic Research, Inc.
- Rodrik D. (2007). *What does the "Rule of Law" mean?* Retrieved December 12, 2007 from: http://rodrik.typepad.com/dani_rodriks_weblog/2007/11/what-does-the-r.html
- Rodrik, D. (2008). *Second best institutions*. Cambridge, MA: Harvard University Press.
- Rodrik, D., Subramanian, A. & Trebbi, F. (2002). *Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development* NBER Working Paper No. 9305. Retrieved July 2008, from <http://ksghome.harvard.edu/~drodrik/institutionsrule,%205.0.pdf>
- Roodman, D. (2005). *An Index of Donor Performance*. Working Paper No. 67, Washington, D.C.: Center for Global Development.
- Sachs J. (2003). *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income*, NBER Working Paper No. 9490.
- Sachs, J. & McArthur, J. W. (2001). *Institutions and Geography: Comment on Acemoglu, Johnson, and Robinson (2000)*. NBER

- Working Paper 8114, Cambridge, MA.
- Sachs, J. & Warner, A. (1995). *Natural Resource Abundance and Economic Growth*. NBER Working Paper 5398.
- Sala-i-Martin, X. (1997). I just ran 2 million regressions. *American Economic Review* 87(2), 178-183.
- Social Sciences & Humanities Inter-divisional Research Ethics Committee, University of Oxford (2008), *Protocol for Elite Interviews*. Retrieved February 2009, from weblearn.ox.ac.uk/site/socsci/appsoc/gradresspsw/31/DREC%20Elite%20Interviewing%20protocol.doc
- Standard and Poor (2003). *Global Stock Markets Factbook*. New York, NY.
- Stinchcombe, A. L. (1965). Social structure and organizations. In J. G. March (Ed.), *Handbook of organizations* (pp. 142-193). Chicago: Rand-McNally.
- Teorell, J., Sören, H. & Bo, R. (2008). *The Quality of Government Dataset. Version 15, May 2008*. University of Gothenburg: The Quality of Government Institute. Retrieved June 2008 from: <http://www.qog.pol.gu.se>
- The Economist. (3 March 2008). *Order in the jungle*. Retrieved April 2008 from http://www.economist.com/displaystory.cfm?story_id=10849115
- Thomas, M. A. (2007). What do the Worldwide Governance Indicators measure? Working Paper. Baltimore, Maryland: Johns Hopkins University.
- Thompson, T. & Shah, A. (2004). Transparency International's Corruption Perceptions Index: Whose Perceptions Are They Anyway? Operations Evaluations Department, Washington, D.C.: World Bank, unpublished.
- Tolbert, P. S. & Zucker, L. G. (1983). Institutional sources of change in the formal structure of organizations: The diffusion of civil service reform, 1880-1935. *Administrative Science Quarterly*, 28(1), 22-39.

- Transparency International (2004). *Frequently Asked Questions about the TI Corruption Perceptions Index (CPI) 2004*. Retrieved June 2006, from http://www.transparency.org/policy_research/surveys_indices/cpi/2004/cpi2004_faq.en.html#five
- UBS (2005), *Handbook 2004/2005*, Zürich. Retrieved July 2008, from www.ubs.com/1/e/investors/annual_reporting2004/handbook.html
- Trebilcock, M. & Daniels, R. (2008). *Rule of Law Reform and Development*. Edward Elgar.
- Tsoukas, H. (1997). The Tyranny of Light: the Temptations and Paradoxes of the Information Society. *Futures*, 29(9), 827-843.
- UNDP (2007). Governance Indicators: A users' guide. Retrieved June 2007, from http://www.undp.org/oslocentre/docs07/undp_users_guide_online_version.pdf
- Van de Walle, S. (2005). *Measuring bureaucratic quality in governance indicators*. Paper for the EGPA Annual Conference, Study Group on Productivity and Quality in the Public Sector, Bern, Switzerland, 31 August - 3 September 2005.
- Van de Walle, S. (2006). The state of the World's bureaucracies. *Journal of Comparative Policy Analysis*, 8(4), 439-450.
- Van de Walle, S. (2008). Comparing the performance of national public sectors: Conceptual problems. *International Journal of Productivity and Performance Management*, 57(4), 329 - 338.
- Van de Walle, S. & Roberts, A. (2008). Publishing performance information: An illusion of control? In W. Van Dooren & S. Van de Walle (Eds.), *Performance information in the public sector: How it is used* (pp. 211-226). Houndmills: Palgrave.
- Van de Walle, S. (forthcoming). International comparisons of public sector performance: how to move ahead? *Public Management Review*.
- Williamson, J. (2000). What Should the World Bank Think About the Washington Consensus? *World Bank Research Observer*, Vol.

15(2), 251-264.

- World Bank (2001). OED IDA Review: Review of the Performance-Based Allocation System, IDA 10-12. February 14, 2001, from wbln0018.
[worldbank.org/oed/oeddoclib.nsf/View+to+Link+WebPages/310529285EC4A83685256A680079BC9F/\\$FILE/PBA.pdf](http://worldbank.org/oed/oeddoclib.nsf/View+to+Link+WebPages/310529285EC4A83685256A680079BC9F/$FILE/PBA.pdf)
- World Bank (2004a). *World Development Report 2005: A Better Investment Climate for Everyone*. Washington, D.C.: World Bank.
- World Bank (2004b). *World Bank Annual Report 2004*. Washington, D.C.: World Bank. Retrieved February 2008, from [http://wbln0018.worldbank.org/oed/oeddoclib.nsf/View+to+Link+WebPages/310529285EC4A83685256A680079BC9F/\\$File/PBA.pdf?OpenElement](http://wbln0018.worldbank.org/oed/oeddoclib.nsf/View+to+Link+WebPages/310529285EC4A83685256A680079BC9F/$File/PBA.pdf?OpenElement)
- World Bank (2004c). Country Policy and Institutional Assessments: 2006 Assessment Questionnaire. Retrieved February 2008, from <http://siteresources.worldbank.org/IDA/Resources/CPIA2006Questionnaire.pdf>
- World Bank (2007). Strengthening World Bank Group Engagement on Governance and Anticorruption. Washington D.C.: World Bank. Retrieved May 2008, from <http://www.worldbank.org/html/extdr/comments/governancefeedback/gacpaper-03212007.pdf>
- World Bank (2005). *World Development Indicators 2005*. Washington, D.C.: World Bank.
- World Bank (2007). *World Development Indicators 2007*. Washington, D.C.: World Bank.
- World Bank (2008a). *World Development Indicators 2008*. Washington D.C.: World Bank.
- World Bank (2008b). Governance, Growth, and Development Decision-making: Reflections by Douglass North, Daron Acemoglu, Francis Fukuyama and Dani Rodrik.
- World Bank Institute (2006). Governance Data: Web-Interactive Inventory of Datasets and Empirical Tools. Retrieved February

- 2008, from
www.worldbank.org/wbi/governance/govdatasets/index.html
- World Bank Institute (2007). *Consolidated Matrix of External and Internal World Bank Governance Data Sets and Instruments*. Washington, D.C.: World Bank Institute. Retrieved February 2008, from
http://siteresources.worldbank.org/EXTWBIGOVANTCOR/Resources/Matrix_2.xls
- World Bank Institute (2008). WGI Press Releases and Media Coverage: Governance Matters VI. Retrieved May 2008 from
<http://web.worldbank.org/WBSITE/EXTERNAL/WBI/EXTWBIGOVANTCOR/0,,contentMDK:21407860~pagePK:64168445~piPK:64168309~theSitePK:1740530,00.html>
- World Economic Forum (2006). *Global Competitiveness Report 2006-2007*. Hampshire: Palgrave Macmillan.
- Zak, P. J. & Knack, S. (2001). Trust and Growth. *Economic Journal*, 111(470), 295-321.
- Zucker, L. G. (1977). The role of institutionalization in cultural persistence. *American Sociological Review*, 42(5), 726-743.

Appendix: Interview methodology

I. Introduction

This section describes the methodology used for the interviews referred to in chapter three (“Uses of governance indicators”) and five (“Politics of governance ratings”). It starts with a discussion of general issues and how they were dealt with. It ends with detailed information about each of the five groups of interviewees: i) investors, ii) donors, iii) World Bank staff, iv) other international organizations, and v) academics.

Interviewees were selected on the basis of their professional position (e.g. Heads of Risk Management, because they have knowledge of the use of indicators in Risk Management). They were usually contacted by e-mail, detailing the nature of the research and the types of information that would be asked from them. In line with standard elite interviewing practices, confidentiality and anonymity was assured to respondents (Dexter, 2006). Most interviews were conducted over the phone. Information was stored so that only the author(s) had access to it. The use of direct quotes and contextual information was checked with interviewees (Social Sciences & Humanities Research Ethics Committee, University of Oxford, 2008). Where possible, views were reported in a more aggregated way (Social Sciences & Humanities Research Ethics Committee, University of Oxford, 4). While more specific information and examples would have enriched the text, such an approach would have compromised the confidentiality and anonymity agreement. Unless

explicit permission was given by an interviewee, he/she (and in some cases the corresponding organization) should not be identifiable.

Most interviews were semi-structured, relying in most cases on open-ended questions and allowing the respondents to engage in wide-ranging statements (see each group of interviewees for details below). By this means, the interviewing technique aimed at providing “opportunity for respondents to organize their answers within their own frameworks” (Aberbach and Rockman, 2002, p. 674). Simultaneously, a systematic approach was applied to minimize the risks of validity and reliability problems conditional on the valuable flexibility of open-ended questions (Berry, 2002).

The existence of sampling bias in the selection of interviewees cannot be excluded. It results mainly from the difficulty, if not impossibility in some cases, to get access to i) interviewees and to ii) the information on the population’s characteristics that would be necessary to draw a representative sample. Bias resulting from non-responses (Goldstein, 2002) cannot be excluded. In particular, access without recommendations was very difficult. Care was taken to rely on multiple sources (Berry, 2002) with different characteristics and to ensure that interviewers with potentially different opinions were selected. As sampling bias cannot be excluded and statistical representativity of interviewees cannot be guaranteed, no hypothesis tests were undertaken and no claims were made about statistical reliability.

I.1 Interviews of corporate and financial investors

The following persons were interviewed in their function as representatives of corporate and financial investors:

1. Senior Manager, France's *Caisse de Dépôts et Consignations*
2. Manager of a leading multinational corporation in the concession of water and sanitation infrastructure in developing countries
3. Head of Risk Management, major international financial investor
4. Head of Risk Management, major international financial investor
5. Head of Risk Management, major international financial investor
6. Head of Risk and Competition Analysis, major international corporate investor
7. Head of Risk and Competition Analysis, major international corporate investor
8. Head of Risk Management, major international corporate investor
9. Analyst, in charge of construction of risk ratings, major international financial investor
10. Analyst, major international financial investor

These interviews were conducted in 2005. In addition, interviews were conducted with John D. Sullivan, Executive Director, Center for International Private Enterprise, Washington D.C. and with a representative of the Institute of International Finance (IIF).

Enclosed to the e-mails that were sent to contact interviewees was the following attachment:

In the framework of an OECD Development Centre Research project on governance indicators, we are interested in understanding to what extent international banks and companies use governance indicators and for what purpose they use these indicators.

Some examples of governance indicators are the six composite governance indicators from the World Bank Institute (Voice and Accountability, Political Stability, Rule of Law, Government Effectiveness, Regulatory Quality and Control of Corruption), the International Country Risk Guide, the Growth Competitiveness Index of the World Economic Forum, Political Rights and Civil Liberties from Freedom House, Pricewaterhouse Cooper's Opacity Index, Latinobarometro and the East Asia Barometer, the Index of Economic Freedom by the Heritage Foundation and the Wall Street Journal, and Transparency International's Corruption Perceptions Index.

1. Does your company/bank use the governance indicators of Kaufmann et al. from the World Bank Institute (Voice and Accountability, Political Stability, Rule of Law, Government Effectiveness, Regulatory Quality and Control of Corruption, <http://www.worldbank.org/wbi/governance/pubs/govmatters3.html>)
2. If yes, for what purpose do they use these indicators?
3. What other governance indicators does your company/bank use for what specific purpose?
4. If your company/bank does not use any such governance indicators, why not?

Other questions asked during the interview aimed at understanding how corporate and financial investors perceive the respective strengths and weaknesses of indicators for the purposes

they use them for. Interviews also aimed at finding out to what extent investors are aware of the methodology and assumptions behind the indicators they use. Corporate and financial investors are more used to the term “Corporate Governance” than to “Public Governance”. In order to avoid misunderstandings, it needed to be clarified in advance that interviews were about Public Governance indicators and not Corporate Governance indicators.

I.2 Bilateral donors

Interviews in the field of bilateral donors were conducted with six spokespersons of international development agencies/ ministries in charge of development from six different countries. These interviews were conducted in 2005 and 2006 and complemented by official documents⁸¹.

Open-ended questions aimed at understanding i) what governance indicators were used for what purpose, ii) if the WGIs were used and if yes, for what purpose, iii) if users of WGIs saw limitations and problems in using them and, if yes, iv) why they still used the WGIs.

⁸¹ After this PhD thesis was completed, the OECD made a detailed overview of governance assessments conducted by donors publicly available (see www.oecd.org/dac/governance/govassessment).

I.3 World Bank staff

From 2007 to 2008, I conducted twelve informal interviews with World Bank senior staff members. More than half of the interviewees either worked in World Bank country offices or had just recently come back from the field. Care was taken to capture different perspectives. To be selected for the interview, staff members had to have a high level of responsibility within the organization and knowledge of the use of governance indicators in the World Bank.

Interviewees were asked about their experience with World Bank governance indicators such as the WGI (formerly KKZ), the CPIA or the Doing Business indicators. Some examples of further questions asked are:

- Did these indicators help to improve local governance?
- Were there controversies around them? If yes, please explain.
- Do you think that these indicators are helpful for your work?

Interviews also aimed at complementing official information on World Bank use of indicators, to find out to which extent World Bank Indicators are used in practice.

I.4 Staff of other international organizations

Five members of other international organizations, whose work is related to governance indicators, were interviewed:

1. OECD official (2005)
2. OECD official (2008)

3. OECD official (2008)
4. UNDP official (2005)
5. UNDP official (2008)

No person was interviewed twice. Informal and unstructured interviews aimed at understanding their view on i) the usefulness of different types of governance indicators, ii) the motivation of different groups to use governance indicators, iii) future trends.

I.5 Academics

A literature search for articles using governance indicators resulted in more than one hundred hits (see for instance Aron, 2000 for an overview). The interviews in the academic field were conducted with two categories of researchers: researchers using indicators (5) and researchers doing research on indicators (3). Attention was paid to a balance in choosing interviewees, so that both critical and less critical voices were represented. Three of the interviews were conducted in 2005, five in 2007.

Open questions aimed at understanding i) the motivation for choosing specific indicators, ii) the depth of knowledge about methodology, strengths and limitations, iii) the level of awareness of alternatives and iv) the reasons for choosing specific indicators despite awareness of limitations and consequences for reliability of research findings.

Samenvatting

Dit proefschrift bevat een uitgebreide analyse van de problemen rond de constructie en het gebruik van indicatoren voor de kwaliteit van het openbare bestuur, bekend als “governance” indicatoren. De kwaliteit van het bestuur in ontwikkelingslanden en opkomende markteconomieën is de laatste jaren onder de aandacht gekomen van internationale investeerders en officiële organisaties voor ontwikkelingssamenwerking in het kader van de OESO, zowel op nationaal als op multilateraal niveau, om een aantal redenen: i) de spectaculaire stijging van de internationale investeringen in ontwikkelingslanden; ii) het einde van de Koude Oorlog; iii) de mislukte hervormingen van het ontwikkelingsbeleid in de jaren 1980 en 1990; en iv) een opkomend nieuw bewustzijn van het belang van de politiek in de economische ontwikkeling en beleidshervormingen.

Als gevolg van de groeiende belangstelling voor het begrip “goed bestuur” in de afgelopen tien jaar door internationale investeerders, hulpdonoren en ontwikkelingsanalisten, stijgt het aanbod van governance indicatoren. Tenminste 140 verschillende stellen van governance indicatoren zijn tegenwoordig beschikbaar, maar een veel kleiner aantal worden intensief gebruikt. Sommige van de meest populaire indicatoren zijn reeds ontstaan voor de recente golf van belangstelling voor het onderwerp, of zijn gebaseerd op gegevens die reeds eerder werden geproduceerd.

Gebruikersgroepen – zoals internationale investeerders, journalisten, officiële nationale en multilaterale hulporganisaties,

ontwikkelingsanalisten en academici (dit zijn meestal mensen die buiten de ontwikkelingslanden leven) – hebben de neiging governance indicatoren te gebruiken – en vaak ook te misbruiken – in hun besluitvormingsprocessen om de kwaliteit van het openbare bestuur te vergelijken, zowel tussen landen onderling als over de tijd heen.

De Wereldbank speelt een leidende rol in de ontwikkeling en productie van internationaal vergelijkbare governance indicatoren op basis van zowel geaggregeerde metingen (bijvoorbeeld de WGIs), als bedrijfsenquêtes (bv. BEEPs) en recent meer actiegerichte indicatoren (bijvoorbeeld PEFA) . De Wereldbank heeft meerdere redenen om de governance indicatoren en afgeleide ratings te produceren. Ten eerste, ratings kunnen (in principe) helpen om hervormingsgebieden te identificeren en het succes van bestuurlijke hervormingen te meten. Ten tweede, governance indicatoren helpen om de kwaliteit van het openbare bestuur op de agenda te zetten, en stimuleren ontwikkelingslanden om hun bestuur te verbeteren. Ten derde, de governance indicatoren kunnen de dialoog tussen landen aanwakkeren, althans indien ze op de juiste wijze worden gebruikt. Ten vierde, de indicatoren kunnen bijdragen tot een grotere transparantie van de beslissingen bij de toewijzing van ontwikkelingshulp. Ten vijfde, indicatoren zijn bruikbaar voor kwantitatieve analyse. Ten zesde, er zijn ook een aantal institutionele motieven voor de productie van governance indicatoren.

De Worldwide Governance Indicators (WGI) worden op grote schaal gebruikt, zowel binnen als buiten de Wereldbank, om

vergelijkingen en rankings te maken in termen van goed bestuur, over de tijd en tussen landen, als randvoorwaarde bij de toewijzing van ontwikkelingshulp, voor risico-evaluaties, voor academische analyses en als informatiebron voor berichten in de media. Echter, er zijn minstens vijf onderling samenhangende problemen met deze indicatoren, waar de gebruikers doorgaans geen rekening mee houden: i) de waarschijnlijke correlatie van storingen, welke een vertekening van de schattingen als gevolg heeft en waardoor de consistentie van de resultaten in het gedrang komt; ii) het gebrek aan transparantie; iii) het gebrek aan een sterk conceptueel kader; iv) steekproefvertekening; en v) het gebrek aan intertemporele vergelijkbaarheid van de indicatoren.

Het verwaarlozen van de beperkingen van de WGIs heeft ernstige gevolgen voor de beleidsvorming en voor academisch onderzoek. Ten eerste, de toewijzing van middelen die gebaseerd is op de WGI ratings zonder rekening te houden met hun slechte vergelijkbaarheid tussen landen en over de tijd is niet alleen onefficiënt, maar ook onrechtvaardig. Ten tweede, de indicatoren kunnen aanleiding geven tot onnodige en zelfs schadelijke bestuurshervormingen in ontwikkelingslanden. Ten derde, regeringen worden in sommige gevallen (gebrek aan) resultaten verweten waar ze niet verantwoordelijk voor kunnen zijn.

Vanwege de beperkingen van de eerste generaties van indicatoren en om te voldoen aan de hoge verwachtingen, wordt bij de Wereldbank momenteel gewerkt aan het ontwerp van meer "actiegerichte" indicatoren, dat wil zeggen indicatoren die aangeven

welke specifieke hervormingen zouden leiden tot een verbetering van de ratings en die tevens de vooruitgang over de tijd kunnen meten. Dit werk vordert traag, want er moet eerst overeengekomen worden welke de prioritaire actieterreinen moeten zijn. De discussie over de actiegerichtte indicatoren en over de beperkingen van de huidige meest vooraanstaande indicatoren ligt aan de basis van een diepgaand debat rond de aspecten van openbaar bestuur die het belangrijkste zijn voor ontwikkelingslanden in verschillende stadia van ontwikkeling. De controverse over het gebruik van de bestaande indicatoren en de wenselijke toekomstige ontwikkelingen zullen een grote impact uitoefenen op het beleid van de Wereldbank in de ontwikkelingslanden, niet het minst omdat indicatoren een politieke rol spelen in de toewijzing van ontwikkelingshulp en in het bepalen van door te voeren hervormingen.

We onderzoeken of een statistisch robuuste meting mogelijk is van het effect van een van de belangrijkste aspecten van het bestuur, het bestaan van de rechtsstaat of “rule of law”, op de economische groei, waarbij een batterij van methoden voor panel data wordt aangewend. Het onderzoek maakt gebruik van een wereldwijde en veelzijdige verzameling van gegevens, waaronder verscheidene “rule of law” indicatoren die reeds enige tijd (tot maximaal vier decennia) waargenomen kunnen worden. De conventionele bevindingen van de positieve effecten van de rechtsstaat op de economische groei zijn echter in statistisch opzicht niet robuust. De resultaten variëren sterk in functie van de gemaakte keuzen met betrekking tot de gebruikte

indicatoren, de schattingsmethodiek, de steekproefperiode en zelfs de schaling van de “rule of law” indicator.

Dit onderzoek leidt tot een aantal aanbevelingen voor beleidsmakers en ontwikkelingsonderzoekers. Op de korte termijn moet rekening gehouden worden met een grote diversiteit aan gegevens die openbaar toegankelijk zijn en gedocumenteerd worden in betreffende handleidingen en inventarissen. Hierbij dienen producenten de informatie te verbeteren waarmee ze hun indicatoren wereldkundig maken en de transparantie van hun werkwijze te vergroten, om de kans op misverstanden omtrent de betekenis, de vergelijkbaarheid en de robuustheid van de indicatoren te beperken.

Op de middellange tot lange termijn zullen meer specifieke, zinvolle en actiegerichte indicatoren worden ontwikkeld. Opdat deze indicatoren daadwerkelijk een toegevoegde waarde hebben voor de ontwikkelingslanden zelf zijn transparantie en actiegerichtheid echter nog niet voldoende. Even cruciaal is het te weten welke concrete acties, welke specifieke bestuurlijke hervormingen de meeste kans hebben om te leiden tot een effectieve vermindering van de armoede en tot duurzame economische groei, en dientengevolge de hoogste prioriteit moeten krijgen. Tegenwoordig maken talrijke ontwikkelingsonderzoekers en praktijkmedewerkers zich zorgen over het feit dat het "good governance" agenda overbelast is en dat een bevolkingsbreed groeiproces heel andere instellingen vereist dan deze welke gepromoveerd worden in de meest gebruikte indicatoren. De vragen die opgeroepen worden en de conflicten die ontstaan bij het

ontwerpen van nieuwe, transparante, actiegerichte indicatoren vormen een aanloop, samen met meer kwalitatief onderzoek, voor het debat betreffende de bestuursprincipes die moeten bevorderd worden om duurzame ontwikkeling te stimuleren, afhankelijk van de lokale context. Het is veel beter om dit debat de dag van vandaag te voeren vooraleer men op grote schaal aanvangt met de uitvoering van bestuurlijke hervormingen die meer onbedoelde groeibelemmeringen zouden kunnen in het leven roepen dan dat zij een breed en duurzaam groeiproces in gang helpen zetten. Menig eerdere poging om de ontwikkeling aan te zwengelen is, alle goede bedoelingen ten spijt, op soortgelijke manier gestrand. Meer robuuste en transparante governance indicatoren zouden een belangrijke bijdrage betekenen voor het onderzoek naar de verbanden tussen goed bestuur en duurzame groei in ontwikkelingslanden. Zij zouden kunnen helpen sequentieregels te identificeren en synergieën uit te baten die cruciaal kunnen zijn voor het succes van specifieke hervormingen van goed bestuur in specifieke domeinen.

Curriculum Vitae Christiane Arndt

CURRICULUM VITAE

Christiane Elisabeth Arndt

PROFESSIONAL EXPERIENCE

Since November 2008 Policy Analyst/Economist	OECD , Public Governance and Territorial Development Directorate, Regulatory Policy Division, Paris, France
March – April 2008 Consultant	World Bank , Public Sector Group, Washington D.C., United States
September – November 2007 Consultant	World Bank , Public Sector Group, Washington D.C., United States
February – December 2007 Research Fellow	Harvard University , John F. Kennedy School of Government, Wiener Center for Social Policy, Cambridge, MA, United States
October 2004 – December 2006 Research Associate	OECD Development Centre , Paris, France
October 2004 – November 2008 Research Fellow/Project Manager	Maastricht Graduate School of Governance , Maastricht, the Netherlands
January – April 2002 Research Assistant	Bertelsmann Foundation , Center for Higher Education Development, Gütersloh, Germany

EDUCATION

October 2004 – June 2009 Ph.D. in Political Economy	Maastricht University, Graduate School of Governance Maastricht, the Netherlands
September 2002 – July 2004 Diplôme de Sciences Po	Institut d'Etudes Politiques de Paris (Sciences Po) Paris, France
September 1999 – March 2004 M.A. Int. Business Studies	Maastricht University, Faculty of Economics and Business Administration Maastricht, the Netherlands

AWARDS

January 2000 – March 2004 National Scholarship	German National Merit Foundation,
--	--

PUBLICATIONS

Arndt, Christiane (2008). The Politics of Governance Ratings. *International Public Management Journal*, 11(3), 275-297.

Oman, Charles & Arndt, Christiane (2008). "La mesure de la gouvernance" in *La gouvernance démocratique : Un nouveau paradigme pour le développement ?* Paris : Karthala.

Oman, Charles & Arndt, Christiane (2007). "Governance Indicators" in Mark Bevir (ed.). *Encyclopedia of Governance*. Sage Publications.

Arndt, Christiane & Oman, Charles (2006). "Korruption und gute Regierungsführung-was wissen wir wirklich?". OECD Development Centre Policy Insight No. 32. Paris: OECD Development Centre.

Oman, Charles & Arndt, Christiane (2006). "Governance Indicators for Development". OECD Development Centre Policy Insight No. 33. OECD Development Centre, Paris.

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

WORK IN PROGRESS

Arndt, Christiane & Crombrugghe, Denis. “The Rule of Law and Economic Growth: A panel data re-examination”. Maastricht Graduate School of Governance, Maastricht. Unpublished manuscript.

Dixon, Ruth, Müllers, Manuel, Arndt, Christiane & Hood, Christopher, “Where International Rankings Meet Negativity Bias: Domestic Press Responses to International Educational Rankings in Three EU Countries”. Oxford University, Oxford and Maastricht Graduate School of Governance, Maastricht. Unpublished manuscript.

Arndt, Christiane & Romero, Carmen (2008). “Review of the policy utility of the World Governance Indicators 2007 for the Central American Countries”, Working Paper Series on Public Sector Management. World Bank, Washington D.C. Unpublished manuscript.

PRESENTATIONS

Invited speaker at

- Rotterdam School of Management, Erasmus University seminar on “Uses and Abuses of Governance Indicators“ in Rotterdam, the Netherlands (2008)
- World Bank, Office of the Chief Economist and Senior Vice President Development Economics, seminar on “The Empirics of Governance” in Washington D.C., USA (2008)
- OECD conference on “Governance Assessments and Aid Effectiveness” in London, U.K. (2008)
- Maxwell School of Syracuse University Seminar on the “Politics of Governance Ratings” in Syracuse, USA (2007)
- IPMN Workshop on “Ranking and Rating Public Services” at Worcester College, Oxford, Great Britain (2007)
- World Bank, Independent Evaluation Group, Seminar on “Uses and Abuses of Governance Indicators”, Washington D.C., USA (2007)
- Harvard University Seminar on “Measuring Governance and Media Capture” in Cambridge, MA, USA (2007)
- DIE Seminar on “Uses and Abuses of Governance Indicators” in Bonn, Germany (2006)
- BIDS Workshop on “Governance and Development”, Dhaka,

Bangladesh (2006)

- Maastricht Graduate School of Governance Seminar on “The Validity of Governance Indicators”, Maastricht, the Netherlands (2006)
- INERPOST Workshop on “Governance, Transparency, and Development”, Athens, Greece (2005)