Governance Matters IV:
Governance Indicators for 1996-2004

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi

The World Bank

May 2005

Abstract: This paper presents the latest update of our aggregate governance indicators, together with new analysis of several issues related to the use of these measures. The governance indicators measure the following six dimensions of governance: i) Voice and Accountability; ii) Political Instability and Violence; iii) Government Effectiveness; iv) Regulatory Quality; v) Rule of Law, and, vi) Control of Corruption. They cover 209 countries and territories for 1996, 1998, 2000, 2002, and 2004. They are based on several hundred individual variables measuring perceptions of governance, drawn from 37 separate data sources constructed by 31 different organizations. We present estimates of the six dimensions of governance for each period, as well as margins of error capturing the range of likely values for each country. These margins of error are not unique to perceptions-based measures of governance, but are an important feature of all efforts to measure governance, including objective indicators. In fact, we provide examples of how individual objective measures provide an incomplete picture of even the quite particular dimensions of governance that they are intended to measure.

We also analyze in some detail changes over time in our estimates of governance; provide a framework for assessing the statistical significance of changes in governance; and suggest a simple rule of thumb for identifying statistically significant changes in country governance over time. The ability to identify significant changes in governance over time is much higher for our aggregate indicators than for any individual indicator. While we find that the quality of governance in a number of countries has changed significantly (in both directions), we also provide evidence suggesting that there are no trends, for better or worse, in global averages of governance. Finally, we interpret the strong observed correlation between income and governance, and argue against recent efforts to apply a discount to governance performance in low income countries.

The data, as well as a web-based graphical interface, are available at: www.worldbank.org/wbi/governance/govdata/. The Appendices and a synthesis of the paper are available at: www.worldbank.org/wbi/governance/pubs/govmatters4.html.

1818 H Street NW, Washington, D.C. 20433. dkaufmann@worldbank.org, akraay@worldbank.org, mmastruzzi@worldbank.org. The opinions expressed here, as well as the governance research indicators themselves, do not reflect the official views of the World Bank, its Executive Directors, or the countries they represent. We would like to thank S. Radelet for excellent feedback, and M. Levy, G. Dunn, A. Karatnycky, R. Fullenbaum, A. Williamson, A. Bellver, S. Weber, D. Cingranelli, D. Richards, R. Writer, M. Wolkers, C. McLiesh, M. Gibney, C. MacCormac, M. Seligson, E. Kite, E. Hart, T. Sealy, D. West, M. Carballo, F. Ndukwe, M. Lagos, A. Lopes-Claros, R. Coutinho, S. Mannan, and D. Cieslikowsky for providing data and answering our numerous questions. The support and collaboration of the World Economic Forum, the U.S. State Department, and the Netherlands Government is appreciated.

1. Introduction

This paper presents the latest update of our aggregate governance indicators, together with new results on the relative importance of perceptions-based and objective indicators; the significance of measured changes over time in governance; and the role of per capita income in cross-country governance comparisons. The governance indicators measure the following six dimensions of governance: i) Voice and Accountability; ii) Political Instability and Violence; iii) Government Effectiveness; iv) Regulatory Quality; v) Rule of Law, and, vi) Control of Corruption. They cover 209 countries and territories for 1996, 1998, 2000, 2002, and 2004. The indicators are based on several hundred individual variables measuring perceptions of governance, drawn from 37 separate data sources constructed by 31 different organizations. We assign these individual measures of governance to categories capturing key dimensions of governance, and use an unobserved components model to construct six aggregate governance indicators in each period. We present the point estimates of the dimensions of governance as well as the margins of errors for each country and period.

We begin by describing the data used to construct this round of the governance indicators in Section 2 of this paper. As discussed in more detail below, we have incorporated information from a substantial number of new data sources, relative to our last set of indicators for the period 1996-2002. Since some of these data sources are also available in earlier periods, we have updated our governance estimates for this earlier period as well. As we have emphasized in our previous work, an attractive feature of the aggregation method we use is that it provides us with not only estimates of governance for each country, but also with measures of the precision or reliability of these estimates, for every country, indicator, and year. The addition of data has improved the precision of our governance indicators relative to previous years. However, the margins of error associated with estimates of governance are not trivial. This implies that cross-country comparisons of levels of governance should continue to be made with due caution. We also underscore that these margins of error are not unique to perceptions-based measures of governance, but are an important feature of all efforts to measure governance, including objective indicators.

Reformers in many governments as well as civil society and investors increasingly view governance as key for development and the investment climate, which in turn has increased the demand for monitoring the quality of governance in a country over time. Further, aid donors have also come to the view that aid flows have a stronger impact on development in countries with good institutional quality, and thus increasingly utilize measurable performance indicators –within which governance features prominently-- for monitoring, evaluation and decision-making at a country level.[1] In light of this, it is also important to measure and interpret trends over time in governance. This we address in Section 3 of the paper, where we discuss how the inevitable measurement error in both subjective and objective indicators of governance affects the conclusions that can be drawn from observed changes over time in such measures.

The most basic insight is that measurement error should temper the conclusions about actual changes in governance based on changes in any individual indicator, while aggregate indicators such as those we develop here can be more informative about changes over time in governance. In addition to this basic insight, we highlight two opposing forces that affect the interpretation of changes over time. On the one hand, if governance itself changes very slowly over time, then observed changes in the data will overstate the magnitude of actual changes in governance. On the other hand, if measurement error is also very persistent over time, then observed changes in the data will understate changes in governance. By providing a framework for assessing the statistical significance of changes in governance over time, we show how these key parameters can be estimated and argue that the former effect dominates, suggesting that changes over time in the governance indicators should be interpreted with some caution. We suggest a simple rule of thumb for identifying statistically significant changes in country governance over time, and find that governance in a number of countries has either significantly improved or deteriorate over the relative short eight-year time span covered by our data. We also document that there is little evidence of any trends – for better or worse – in global averages of governance.

[1] For example, the International Development Association (the highly concessional loan window of the World Bank) relies heavily on the World Bank's Country Policy and Institutional Assessment, one of the ingredients in our aggregate governance indicators. The U.S. government's Millennium Challenge Account bases country eligibility in part on five of our governance indicators.

The margins of error we emphasize are not unique to the perceptions data we use to construct our aggregate governance indicators: measurement error is pervasive among all measures of governance and institutional quality. An advantage of our measures of governance is that we are able to be explicit about the accompanying margins of error, whereas these are most often left implicit with objective measures of governance. In Section 4 of this paper we investigate in more detail discrepancies between subjective and objective measures of very specific dimensions of the regulatory environment. We show that firms' survey responses about their tax burden, and the ease of starting a new business, reflect not only the *de jure* regulations governing these issues, but also the overall institutional and governance environment in which these regulations are applied. This finding emphasizes the importance of relying on a full range of measures of governance, and not exclusively subjective or objective measures, when assessing the quality of governance across countries.

In the final section of the paper we turn to two issues that arise when interpreting the strong positive correlation observed between measures of governance and per capita incomes. One critique of subjective or perceptions-based governance measures is that they are subject to "halo effects" – respondents rating countries might provide good governance scores to richer countries simply because they are richer. While this is certainly a possible source of bias, we show that it will lead to a significant upward bias in the correlation between income and governance only if these halo effects are implausibly strong. The second issue concerns the interpretation of the quality of governance in low income countries, with particular application to Sub-Saharan Africa, where the international community is rightly focusing its attention in the effort to meet the Millennium Development Goals of halving poverty by 2015. Although countries in the region on average tend to score quite poorly on most measures of governance, some observers have argued that this poor governance performance should be discounted because per capita incomes in the region are also low. Implicit in this argument is the view that there is a strong causal impact of incomes on governance. However, we argue that existing evidence does not support a strong causal channel operating in this direction – most of the correlation between governance and per capita incomes reflects causation from the former to the latter. In light of this we suggest that it would be inappropriate to divert attention from the weak average governance performance of the region (while also recognizing the individual countries that are strong governance

performers in the region), simply because the region is poor. While we focus on Africa because of the recent emphasis in the aid community on the region, the fallacy of discounting the extent of misgovernance in a country or region due to low incomes applies more generally to any setting with poor governance and low incomes.

We conclude by summarizing the key findings and noting the policy implications of our work

2. Updated Governance Indicators for 1996-2004

In this section we briefly describe the update of our governance indicators for 2004, as well as some minor backwards revisions to the indicators for 1996-2002. Our basic methodology has not changed from past years, and a detailed discussion can be found in Kaufmann, Kraay, and Mastruzzi (2004). We construct measures of six dimensions governance:

1. *Voice and Accountability* – measuring political, civil and human rights
2. *Political Instability and Violence* – measuring the likelihood of violent threats to, or changes in, government, including terrorism
3. *Government Effectiveness* – measuring the competence of the bureaucracy and the quality of public service delivery
4. *Regulatory Burden* – measuring the incidence of market-unfriendly policies
5. *Rule of Law* – measuring the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence
6. *Control of Corruption* – measuring the exercise of public power for private gain, including both petty and grand corruption and state capture

In Appendix D we define these six dimensions of governance in more detail.

2.1 Data and Methodology

As in past years we rely on a large number of individual data sources which provide us with information on perceptions of governance. These data sources consist

of surveys of firms and individuals, as well as the assessments of commercial risk rating agencies, non-governmental organizations, and a number of multilateral aid agencies. A full list of these sources is presented in Table 1. For this round of the data, we rely on a total of 352 individual variables measuring different dimensions of governance. These are taken from 37 different sources, produced by 31 different organizations. Appendices A and B provide a detailed description of each data source, and document how we have assigned individual data sources to our six aggregate indicators.

These 37 sources include 12 new data sources for 2004, indicated with asterisks in Table 1. The new sources come from a diverse set of organizations. Three of these come from international organizations, in the form of country assessments prepared by economists at the African Development Bank, the Asian Development Bank, and the United Nations Economic Commission for Africa. Another three are from commercial consultancies: IJET Travel Consultancies, Merchant International Group, and Political and Economic Risk Consultancy.[2] The remaining six come from a mix of NGOs and universities: Bertelsmann Foundation, Brown University Center for Public Policy, the Countries at the Crossroads publication of Freedom House, Fundar, the International Research and Exchanges Board, and Vanderbilt University.[3] Several of these new sources also have data available prior to 2004. In order to make full use of this additional data, as well as to improve the comparability of the governance indicators over time, we have revised our previous indicators for all periods to incorporate these sources. Typically the addition of these sources has very little effect on our past indicators, but it does make them more comparable over time.

It is also important to note that our data sources reflect the perceptions of a very diverse group of respondents. Several of our data sources are surveys of individuals or domestic firms with first-hand knowledge of the governance situation in the country. These include the World Economic Forum's Global Competitiveness Report, the Institute for Management Development's World Competitiveness Yearbook, the World Bank's

[2] The last of these, Political and Economic Risk Consultancy, is not quite a "new" source as it appeared in our 1998 and 2000 indicators in the past, but not in the 2002 and 2004 indicators.
[3] It is worth noting that we do not use the Transparency International Corruption Perceptions Index (CPI) as a component of our aggregate corruption indicator. This is because the CPI is itself an aggregate of a number of individual sources, all of which we have already included in our corruption indicator.

business environment surveys, and a variety of global polls of individuals conducted by Gallup, Latinobarometro, and Afrobarometro. We also capture the perceptions of country analysts at the major multilateral development agencies (the European Bank for Reconstruction and Development, the African Development Bank, the Asian Development Bank, the UN Economic Commission for Africa, and the World Bank), reflecting these individuals' in-depth experience working on the countries they assess. Other data sources from NGOs (such as Amnesty International, Reporters Without Borders, and Freedom House), as well as commercial risk rating agencies (such as EIU and DRI) base their assessments on a global network of correspondents typically living in the country they are rating.

As in our past work, we combine the many individual data sources into six aggregate governance indicators. The premise underlying this statistical approach should not be too controversial – each of the individual data sources we have provides an imperfect signal of some deep underlying notion of governance that is difficult to observe directly. This means that as users of the individual sources, we face a signal-extraction problem – how do we isolate the informative signal about governance from each individual data source, and how do we optimally combine the many data sources to get the best possible signal of governance in a country based on all the available data? In Appendix D we describe in detail the statistical procedure we use to perform this aggregation, known as the unobserved components model. The main advantage of this approach is that the aggregate indicators are more informative about unobserved governance than any individual data source. Moreover, the methodology allows us to be explicit about the precision – or imprecision – of our estimates of governance in each country. As we discuss in more detail throughout the paper, this imprecision is not a consequence of our reliance on subjective or perceptions data on governance – rather imprecision is an issue that should be squarely addressed in all efforts to measure the quality of governance.

2.2 Estimates of Governance 1996 - 2004

In Appendix C we report the aggregate governance indicators, for all countries, for each of the six indicators and for all five periods. The governance estimates are normally distributed with a mean of zero and a standard deviation of one in each period. This implies that virtually all scores lie between -2.5 and 2.5, with higher scores

corresponding to better outcomes.[4] This also implies that our aggregate estimates convey no information about trends in global averages of governance, but they are of course informative about changes in individual countries' relative positions over time. Below we discuss the information in our individual indicators regarding trends over time in global averages of governance.

Table 2 summarizes some of the key features of our governance indicators. In the top panel we show the number of countries included in each of the six indicators and four periods. In 2004 the Government Effectiveness indicator covers the largest set of 209 countries, with the other sources covering between 204 and 208 countries.[5] Over time, there has been a steady increase in the number of sources included in each of our indicators. This increase in the number of data sources is reflected in an increase in the median number of sources available per country, which, depending on the governance component, ranges from four to six in 1996, and from eight to eleven in 2004. Thanks to the increase in sources, the proportion of countries in our sample for which our governance estimates are based on only one source has also declined considerably, to an average of only 7 percent of the sample in 2004.

An important consequence of this expanding data availability is that the margins of error for the governance indicators have declined, as shown in the final panel of Table 2. Depending on the governance component, in 1996 the average (for all countries) of the standard error[6] ranged from 0.26 to 0.36, while in 2004 the corresponding range is from 0.18 to 0.27. These declines in margins of error illustrate the benefits in terms of precision of constructing composite indicators based on as much information as possible. Of course, since our aggregate indicators combine information from all of these sources, they have greater precision than any individual underlying data source. Looking across all five time periods, the median standard error of the individual data

[4] For a handful of cases, individual country ratings can exceed these boundaries when estimates of governance are particularly high or low.
[5] A few of the entities covered by our indicators are not fully independent states (Puerto Rico, Hong Kong, West Bank/Gaza, Martinique, and French Guyana). A handful of very small independent principalities (Monaco, San Marino, and Andorra) are also included. For stylistic convenience all 209 entities are often referred in this paper as "countries".
[6] As described in detail in Appendix D, our outcome of aggregation procedure is a distribution of possible values of governance for a country, conditional on the observed data for that country. The mean of this conditional distribution is our estimate of governance, and we refer to the standard deviation of this conditional distribution as the "standard error" of the governance estimate.

sources for the governance indicators was 0.58, with an interquartile range from 0.45 to 0.84.

Despite this increase in precision as a benefit of aggregation, the margins of error for the aggregate governance indicators are non-trivial. We illustrate this point in Figure 1. In the two panels of Figure 1, we order countries in ascending order according to their point estimates of governance in 2002 on the horizontal axis, and on the vertical axis we plot the estimate of governance and the associated 90% confidence interval described above. We do this for two of the six governance indicators, political stability, and control of corruption. The size of these confidence intervals varies across countries, as different countries appear in different numbers of sources with different levels of precision. The resulting confidence intervals are substantial relative to the units in which governance is measured. To emphasize this point, the horizontal lines in Figure 1 delineate the quartiles of the distribution of governance estimates. Even though the differences between countries in the bottom and top quartiles are substantial, the number of countries that have 90% confidence intervals that lie entirely within a given quartile is not large. From Figure 1 it should also be evident that many of the small differences in estimates of governance across countries are not likely to be statistically significant at reasonable confidence levels. For many applications, instead of merely observing the point estimates, it is therefore more useful to focus on the *range* of possible governance values for each country (as summarized in the 90% confidence intervals shown in Figure 1).

As an illustration of the importance of margins of error in governance comparisons, consider the eligibility criteria for the U.S. Millennium Challenge Account (MCA). Countries' eligibility for grants from the MCA is determined by their relative positions on 16 different measures of country performance. One of these is our Control of Corruption indicator, where countries are required to score above the median among all potentially eligible countries in order to qualify for MCA funding. As we have noted elsewhere, this procedure risks misclassifying countries around the median because the margins of error for such countries often includes the median score. In contrast, for countries near the top and the bottom of potential MCA beneficiaries, we can be quite confident that they do in fact fall above and below the median, respectively.

Table 3 illustrates the role of margins of error in this calculation. We focus attention on the set of 70 countries identified as potential MCA beneficiaries for the 2005 fiscal year.[7] For these countries, we calculate the median score on our Control of Corruption indicator for 2004. Next, using our governance estimates and their accompanying standard errors, for each country we calculate the probability that the country's level of corruption falls above the median for this group. The results of this calculation are summarized in the first column of Table 3. For 17 poorly-performing countries, or about one-quarter of the sample, there is less than a 10 percent chance that corruption in these countries actually falls above the median. For another 23 countries, or about a third of the sample, we are quite confident that corruption in these countries falls above the median, with a probability of at least 90 percent. In contrast, for the remaining 30 countries, the probability that they fall above the median is somewhere between 10 percent and 90 percent, and so we have less confidence that these countries are correctly classified. If we relax our standards of significance to 25 percent and 75 percent, we find that only about 20 countries out of 70, or 29 percent of countries fall in this zone of uncertainty.[8]

This example illustrates the importance of taking margins of error into account when making governance comparisons across countries. Our aggregate governance indicator is able to identify with a fairly substantial degree of confidence groups of countries where the probability that corruption is above or below the median is large. But at the same time there remains an intermediate group of countries where we can be less confident that they are correctly classified as being "good" or "bad" performers based on their point estimates of governance alone.

It is also important to note how this example illustrates the benefit of aggregating many sources of data on corruption, as we do. The remaining columns of Table 3 show perform the same calculations, but relying on successively less precise measures of governance. The second and third columns use our own Control of Corruption

[7] See http://www.mcc.gov/ for details on the MCA eligibility criteria.
[8] We first performed these MCA-related calculations in late 2002, shortly after the announcement of the initial MCA eligibility criteria. At that time, using the older version of our 2000 Control of Corruption indicator, we found that 23 out of 61 countries (or 38 percent of countries) fell in this intermediate zone. This much higher proportion of intermediate countries reflected the fact that the old version of or 2000 Control of Corruption indicator relied on substantially fewer data sources than we now have available to us for both 2000 and 2004.

indicators for 2000 and 1996. These indicators cover fewer countries, and because they rely on a smaller set of sources available at the time, the margins of error for individual countries are higher than in 2004 (see the standard errors reported in the last row). In 1996, for example, 35 percent of the countries for which data is available fall in the intermediate category where the probability that they fall in the top half of the sample is between 25 percent and 75 percent – as opposed to only 29 percent of countries falling in this grey area with the 2004 indicator. The last three columns of the table show the same information for three of our individual sources, WMO, DRI, and GCS. These individual sources have substantially higher margins of error than our aggregate indicators, and in the case of DRI and GCS also cover substantially fewer countries. In addition, we see that there is greater uncertainty about country rankings when relying on just a single indicator: for GCS, for example, the fraction of countries falling in the intermediate category rises to 40 percent. This illustrates the benefit of relying on aggregate indicators which are more informative than individual indicators when trying to classify countries according to their levels of governance.

2.3 Changes over Time in Governance at the Country Level

We now turn to the changes over time in our estimates of governance in individual countries. Figure 2 illustrates these changes for two selected governance indicators over the period 1996-2004. In both panels, we plot the 2004 score on the horizontal axis, and the 1996 score on the vertical axis. We also plot the 45-degree line, so that countries above this line correspond to declines in the quality of governance, while countries below the line correspond to improvements in governance. The first feature of this graph is that most countries are clustered quite close to the 45-degree line, indicating that changes in our estimates of governance in these countries are relatively small over the eight-year period covered by the graph. A similar pattern emerges for the other four dimensions of governance (not shown in Figure 2), and, not surprisingly the correlation between current and lagged estimates of governance is even higher when we consider shorter time periods.

However, our estimates of governance do change substantially for some countries in some periods. For example, from 1996 to 2004, countries like Cote d'Ivoire, Zimbabwe, Nepal and the Central African Republic show substantial declines in, among others, the Voice and Accountability measure, while countries like Argentina and Sierra Leone deteriorate on Regulatory Quality, and Zimbabwe, Cyprus, Israel, and Moldova decline on Control of Corruption measures, contrasting countries like Latvia and Bahrain which show substantial improvements in Control of Corruption, while Croatia, Nigeria, and Bosnia and Herzegovina improve in Voice and Accountability, for instance. [9]

In Figure 2 we have labeled those countries for which the change in estimated governance over the 1996-2004 period is sufficiently large that the 90% confidence intervals for governance in the two periods do not overlap. While this is not a formal test of the statistical significance of changes over time in governance, it is a very simple and transparent rule of thumb for identifying large changes in governance. In the next section of this paper we will discuss in more detail how to assess the statistical significance of changes in governance. We also note that there are of course more "large" changes in governance if we relax our standards to asking whether, say, 75 percent confidence intervals overlap or not. In this case, we would identify an average of 35 large changes per indicator, as opposed to an average of 15 per indicator for non-overlapping 90 percent confidence intervals.

For the rest of this subsection we provide details on why our estimates of governance have changed for those countries where changes are large according to this simple rule of thumb. In Table 4 we provide more detail on all of the large changes in our six governance indicators over the period 1996-2004. The first three columns report the level of governance in the two periods, and the change. The remaining columns provide information on the two main potential sources of changes in our estimates of governance for a particular country: (1) changes over time in individual data sources' assessments

[9] Focusing on the shorter 1998-2004 period (yet one which has a larger country overlap) also yields a number of countries that have undergone large changes, such as the decline exhibited in Control of Corruption, Government Effectiveness and Rule of Law for West Bank/Gaza (for which there was no data in 1996), Ivory Coast, Zimbabwe and Eritrea, and the deterioration in Voice and Accountability during the period in Nepal, Kyrgyz Republic, and Russia, contrasting the improvements in Control of Corruption in the Slovak Republic, Croatia, Serbia, Bulgaria, Madagascar and Colombia, or in Political Stability/Violence in Rwanda, Sierra Leone, Angola, Turkey, South Africa and Senegal, for instance.

of governance, and (2) changes due to the addition of new data sources for a country. Consider first changes over time in the underlying data sources that are available in both periods for a country. In the column labeled "Agree" we report the number of sources available in both periods which move in the same direction as the aggregate indicator. The columns labeled "No Change" and "Disagree" report the number of sources on which that country's score does not change or moves in the opposite direction to the aggregate indicator. For each country we also summarize the extent to which changes in the individual sources agree with the direction of change in the aggregate indicator by calculating the "Agreement Ratio", or "Agree" / ("Agree" + "Disagree").

The agreement ratio is quite high for countries with large changes in governance. Averaging across all countries and indicators, we find an average agreement ratio of 0.86 for the period 1996-2004, as reported in Table 5. For the six indicators the agreement ratio ranges from a low of 0.76 for Government Effectiveness to a high of 0.93 for Voice and Accountability. This provides some confidence that for countries with large changes in our governance estimates, these changes are being driven primarily by changes in underlying sources. In fact, there are only three cases where the agreement ratio is less than one-half: Indonesia and Zambia for Regulatory Quality, and Iceland for Control of Corruption.[10] It is also worth noting that the agreement ratios for large changes in governance are much higher than the agreement ratios for all changes in governance. This can also be seen in Table 5 which computes the same agreement ratio, but for all countries over the period 1996-2004. The agreement ratio averages 64 percent, suggesting that for the more typical smaller changes in our governance estimates, there is much more disagreement across individual sources about the direction of the change than there is for large changes.

The remaining columns of Table 4 measure how the addition of new sources of governance data in 2004 contributes to the change in the estimate of governance for a country. We do this by first calculating what our estimate of governance in the second

[10] For Indonesia, the large decline in the overall score was due to a fairly substantial decline in one underlying source, HER, as well as the addition of new sources in 2004 that provided lower scores than the ones available in both periods. In the case of Iceland, the large improvement seems to be driven entirely by Iceland's big improvement, from an unusually low base, in the score assigned to it by GCS in 1996. Finally, in the case of Zambia the three sources that move in the opposite direction from the aggregate indicator do so only very slightly and these very small improvements are strongly offset by worsening in the remaining two sources.

period would have been had we used only sources available in both periods. We also calculate what our estimate of governance would be if we were to rely only on the new sources added in the second period relative to the first period.[11] If this latter score is higher (lower) than the former, then we know that the new data sources on average rate the country better (worse) than do the existing sources available in both periods, and this effect on its own will contribute to an improvement (decline) in estimated governance for the country. The overall score for the country in the second period is just a weighted average of these two scores. We report these two scores, and the accompanying weights, in the last four columns of Table 4.

Interestingly, and reassuringly, the addition of new sources does not appear to have very substantial effects on the changes over time in the governance estimates. To assess this, we have computed the absolute difference between the "balanced" score and the score based on new sources, and expressed this as a fraction of the absolute change in the overall governance estimate over the two periods. Averaging across all the entries in Table 4 gives a figure of 9 percent. Taken together, this evidence suggests that for the large changes in governance shown in this table, we can have a good deal of confidence that it is mostly driven by changes in the underlying sources on which the aggregate indicators are based. In contrast, we should be much more cautious in our interpretation of many of the smaller changes in our aggregate governance indicators.

2.4 Trends in Global Governance

We now examine the limited available evidence on trends in global averages of governance. As we have already noted, our aggregate governance indicators are not informative about trends in global averages because we have normalized these averages to zero in each period, as a choice of units. While the aggregate indicators are of course informative about the relative performance of individual (or groups of) countries, in order to assess trends in global governance we need to return to our underlying individual data sources.

[11] Of all the large changes we identify in this way, there is only one case where a data source was dropped: Israel, which was dropped by BERI in 2000. For all the remaining cases changes in the composition of data sources reflect only the addition of new sources.

In Table 6 we summarize trends in world averages in a number of our individual data sources. Most of the sources in this table are polls of experts, with data extending over the whole period 1996-2004. Only one of them, GCS, is a survey with sufficiently standard format to enable comparisons over this period of time. The first five columns present the average across all countries of each of the sources in each of the years. The underlying data have been rescaled to run from zero to one, and for each source and governance component, we report the score on the same question or average of questions that we use in the aggregate indicator. The next five columns report the standard deviation across countries for each source. The final column reports the t-statistic associated with a test of the null hypothesis that the world average score is the same in 1996 as in 2004.

The picture that emerges from Table 6 is sobering. There is very little evidence of statistically significant improvements in governance worldwide. The 22 eight-year changes reported here are divided exactly in half into 11 improvements and 11 declines in global averages. Interesting there are nine cases of statistically significant changes at the 10 percent level or better (t-statistics greater than 1.64 in absolute value), and these are split between three improvements and six declines. It is not clear how much importance ought to be ascribed to these trends in world averages. On the one hand, these statistics represent the only information we have on trends over time, and so they should be taken seriously. On the other hand, it is also clear that there is substantial disagreement among sources about even the direction of changes in global averages of governance. For now we cautiously conclude that we certainly do not have any evidence of any significant improvement in governance worldwide, and if anything the evidence is suggestive of a deterioration, at the very least in key dimensions such as regulatory quality, rule of law, and control of corruption.

3. Statistical Significance of Changes in Governance over Time

Reformers in many governments as well as civil society and investors increasingly view governance as key for development and the investment climate, which in turn has increased the demand for monitoring the quality of governance in a country over time. Further, aid donors have also come to the view that aid flows have a stronger impact on development in countries with good institutional quality. In light of this, it is important not only to measure levels, but also to assess trends over time in governance. The presence of measurement error in all types of governance indicators, including our own, makes assessing trends in governance a challenging undertaking. In this section we develop a formal statistical methodology, as well as some simple rules of thumb, for identifying changes in governance that are likely to be statistically and practically significant.

In our description of the data in the previous section we have emphasized the importance of measurement error in governance indicators, and its consequences for interpreting cross-country differences in measures of governance. We have also identified a limited number of episodes in which changes over time in our aggregate governance indicators are large relative to the associated margins of error. In this section of the paper we provide a more formal statistical analysis of changes over time in governance. At a most basic level, it should be clear that the presence of measurement error in the underlying data implies that we should be cautious about reading too much into observed changes in individual and composite measures of governance, both subjective and objective. In this section we formalize this common-sense notion and expand it to consider how persistence over time in both governance and measurement error affect the statistical inferences we can make about changes over time in governance from the available data.

3.1 Changes in Individual Indicators

It is useful to begin our discussion with the simplest possible example of how measurement error impacts our interpretation of changes over time in observed governance indicators, both subjective and objective. Suppose that we have only one source of governance data observed at two points in time, and we want to make

inferences about how governance has changed in a country. To keep notation as simple as possible, we suppress country subscripts and write the observed data at time t, y(t), as the sum of true unobserved governance in that period, g(t), and an error term capturing measurement error:

$$(1) \qquad y(t) = g(t) + \varepsilon(t) \quad , \quad t = 1, 2$$

As a choice of units, we assume that true governance has mean zero and standard deviation one, and that the error term has zero mean. For simplicity we assume that the variance of the error term is the same in both periods and is equal to σ^2. Note that σ^2 is the noise-to-signal ratio in the observed governance data (the ratio of the variance of the error to the variance of unobserved governance). We also allow for the possibility that both governance and the error term are correlated over time, with correlations ρ and r, respectively. Finally we assume that both governance and the error term are normally distributed. With these simplifying assumptions, consider the problem of making inferences about the change in unobserved governance, g(t)-g(t-1), conditional on observing data y(t) and y(t-1) in the two periods. Using the fact that unobserved governance and the data are jointly normally distributed, we can use the properties of the multivariate normal distribution to arrive at the following expressions for the mean and variance of the change in governance, conditional on the observed data:[12]

$$(2) \qquad \begin{aligned} E\left[g(t) - g(t-1) \,|\, y(t), y(t-1)\right] &= \frac{(1-\rho)\cdot\left(y(t) - y(t-1)\right)}{1 + \sigma^2 \cdot (1-r) - \rho} \\ V\left[g(t) - g(t-1) \,|\, y(t), y(t-1)\right] &= \frac{2\cdot(1-\rho)\cdot(1-r)\cdot\sigma^2}{1 + \sigma^2 \cdot (1-r) - \rho} \end{aligned}$$

It is natural to use this conditional mean as our best estimate of the change in governance, and the conditional variance as an indicator of the confidence we have in the estimate. This is in fact exactly analogous to how we obtain estimates of levels of governance and associated standard errors using the unobserved components model described in Appendix D.

[12] The simple example here is a special case of a more general model we discuss below.

To interpret these expressions, consider first the case where there is no persistence in governance or in the error terms, i.e. ρ=r=0. In this case, our estimate of the change in governance is simply $\frac{y(t)-y(t-1)}{1+\sigma^2}$. In particular, we should take the observed change in the single source and scale it down by a factor of $\frac{1}{1+\sigma^2}$ to reflect the fact that the data measures governance with error. It is also clear from Equation (2) that the higher is ρ, the more we should discount observed changes in governance. Intuitively, if we knew that governance changes very slowly over time, then any observed change in the data is more likely to reflect changes in the error term, and so we should discount this observed change more heavily. In the limit where governance is perfectly correlated in the two periods, we would know for sure that any change observed in the data must reflect only fluctuations in the error term, and so we would completely discount the observed change in the data. That is, our estimate of the change in governance would be zero regardless of the observed change in the data.

The effect of persistence in the error terms works in the opposite direction: we should scale down the observed change in the data by less the larger is the correlation over time in the error terms. Again the intuition for this is simple – if we know that the error with which a given source measures governance is persistent over time, then any observed change in the source is likely to understate the true change in unobserved governance. As a result our best estimate of the change in governance will be larger than the observed change in the data. Interestingly, if the correlation in unobserved governance and the error term are equal to each other, i.e. ρ=r, then these two effects offset exactly and the discount applied to the observed change in governance is $\frac{1}{1+\sigma^2}$.

How much confidence should we have in the statistical significance of the change in unobserved governance based on the observed data? Suppose that we observe a change in the indicator equal to k standard deviations of the changes in this variable, i.e. $y(t)-y(t-1)=k\cdot\sqrt{2\cdot\left(1+\sigma^2\cdot(1-r)-\rho\right)}$. Does this signal a significant change in governance? In order to test the null hypothesis that the change in governance is zero, we can construct the usual z-statistic associated with this

hypothesis, i.e. the ratio of the mean of the change in governance conditional on the data to the square root of the conditional variance, which simplifies to:

$$(3) \qquad z = \frac{E\left[g(t) - g(t-1) \mid y(t), y(t-1)\right]}{\sqrt{V\left[g(t) - g(t-1) \mid y(t), y(t-1)\right]}} = \frac{k}{\sigma} \cdot \sqrt{\frac{1-\rho}{1-r}}$$

Not surprisingly, the observed change in the data is more likely to signal a significant change in unobserved governance the larger is the observed change in the data (i.e. the larger is k), and the lower is the signal-to-noise ratio in the data (i.e. the smaller is σ). And building on the intuitions above, the observed change in the data is also more likely to signal a significant change in unobserved governance the lower is the persistence in unobserved governance, ρ, and the higher is the persistence in the error term, r.

Figure 3 puts some numbers to this simple calculation. We graph the number of standard deviations of the observed change in the data, k, on the horizontal axis, and we plot the z-statistic in Equation (3) on the vertical axis for different values of the key parameters. We set σ^2=0.36, as this is the median value for the noise-to-signal ratio across all of the individual data sources we use to construct our six governance indicators in each of the five periods. In an earlier paper we have argued that the noise-to-signal ratio in objective measures of governance is likely to be at least as large as this.[13] The thin upward-sloping line traces out the z-statistic as a function of k for this value of the noise-to-signal ratio, but assuming that the correlation in governance and the error term are zero, i.e. ρ=r=0. The z-statistic is greater than the 90-percent critical value for changes in the observed data that are more than one standard deviation away from the mean change. This suggests that if there is no persistence in governance or in the error terms, quite a large proportion of observed changes in individual governance indicators would in fact signal a significant change in unobserved governance. In fact, if changes in the observed governance indicator are approximately normally distributed, the largest one-third of all absolute changes would signal changes in governance that are significant at the 90% level.

[13] See Kaufmann, Kraay and Mastruzzi (2004)

The bold upward-sloping line corresponds to the more empirically relevant case where there is persistence in both governance and the error terms. The line is drawn for the same noise-to-signal ratio as before, and in addition we assume that the correlation of unobserved governance over time is ρ=0.9 and the correlation in the error term is r=0.4. In the next subsection we show how these parameters can be estimated using our governance data, and find that these values are typical ones. In particular, we shall see shortly that unobserved governance tends to be highly persistent over the eight-year period spanned by our dataset, and although the error terms are also typically positively correlated over time they are much less so than governance. Based on the intuitions developed above, this suggests that much larger observed changes in governance indicators would be required to signal statistically significant changes in unobserved governance. This is exactly what we find. The bold line crosses the 90% critical value at k=2.5, indicating that only those observed changes in the data more than 2.5 standard deviations away from the mean would signal a statistically significant change in governance. Again, if changes in the observed governance indicators are normally distributed, this would imply that only the top one percent of all absolute changes would correspond to significant changes in governance. This in turn suggests that drawing conclusions about changes in governance based on changes in individual governance indicators should be done with an abundance of caution.

In Figure 4 we use *de jure* and *de facto* data on business entry (discussed in more detail in the next section) as an illustration of the difficulty of identifying statistically significant changes over time in governance using individual indicators. In this graph, we plot the change between 2003 and 2004 in the Global Competitiveness Survey question regarding the ease of business entry, against the change in the number of days required to start a business from the Doing Business project of the World Bank (see World Bank (2004)), taken over the same period.[14] We interpret both of these measures as providing noisy signals of changes in the regulatory environment. From the discussion above, only the largest of these changes (in absolute value) are likely to signal statistically significant changes in underlying governance. In particular, if we take our representative assumptions regarding the persistence in governance and in the error terms, we saw that only the top one percent of changes in the observed indicators signal

[14] We would like to thank Caralee McLiesh for kindly providing the unpublished regulation of entry data for 2003.

changes in governance that are significant at the 90 percent level. Even if we relax our standards of significance to 75 percent, only changes in the observed data that are more than 1.8 standard deviations away from the mean, or about the top seven percent of all changes, will signal significant changes in governance in each individual indicator. This translates to roughly five large changes per indicator. We have labeled the top five changes in absolute value for both indicators in Figure 4.

Another striking observation from Figure 4 is that the correlation between the changes over time in these measures is virtually zero. This illustrates the likelihood that relying on individual measures of governance to assess changes over time may lead to very different conclusions depending on which measure is chosen. Further, it also suggests that aggregate indicators which combine information from several different sources might provide a more robust indicator of changes over time in governance. In the next subsection we extend our discussion of the significance of changes over time in governance to the case of composite indicators in order to explore this more fully.

3.2 Changes in Aggregate Indicators

We now elaborate on the previous discussion to address the problem of making inferences about changes over time in country governance based on our aggregate indicators. Just as we found that aggregate indicators are more informative about levels of governance than individual indicators, changes over time in aggregate indicators can be more informative about trends in governance than changes in individual indicators. To formalize this we develop a two-period version of the unobserved components model that we have used to construct the aggregate indicators in each period. We then use it to be more precise about the statistical significance of changes over time in our estimates of governance.

Let y(j,k,t) denote the governance assessment provided by individual data source k in period t for country j. We use a two-period version of the unobserved components model to express this observed data as a linear function of unobserved governance in country j at time t, g(j,t), and an error term capturing the various sources of measurement error that we have been discussing, ε(j,k,t):

(4) $$y(j,k,t) = \alpha(k,t) + \beta(k,t)\cdot\left(g(j,t) + \varepsilon(j,k,t)\right)$$

The intercept and slope parameters $\alpha(k,t)$ and $\beta(k,t)$ vary by data source and over time. As in our single-period model we assume that unobserved governance and the error terms are normally distributed with mean zero. We maintain the identifying assumption that unobserved governance and the all the error terms are mutually independent, i.e. $E\left[g(j,t)\cdot\varepsilon(j,k,s)\right] = 0$ for all sources k and periods t and s, and $E\left[\varepsilon(j,k,t)\cdot\varepsilon(j,m,s)\right] = 0$ for all sources k different from m and for all periods t and s. We also maintain as a choice of units that the variance of unobserved governance is one in each period, i.e. $E\left[g(j,t)^2\right] = 0$ for all t. Our only substantive new assumption is that unobserved governance is correlated over time, as are the error terms, i.e. $E\left[g(j,t)\cdot g(j,t-1)\right] = \rho$, and $E\left[\varepsilon(j,k,t)\cdot\varepsilon(j,k,t-1)\right] = r_k\cdot\sigma(k,t)\cdot\sigma(k,t-1)$, so that ρ and r_k are the correlations over time of governance and the error term in source k, respectively.

Next let y(j,t) denote the Kx1 vector of observed data for each country; $\alpha(t)$, $\beta(t)$, $\sigma(t)^2$ and r denote the Kx1 vectors of the parameters in period t; and let B(t), $\Sigma(t)$ and R denote KxK matrices with the vectors $\beta(t)$, $\sigma(t)^2$ and r on their diagonals. Then using the properties of the multivariate normal distribution, the joint distribution of unobserved governance in the two periods in a country, conditional on the observed data for that country is normal with mean and variance:

(5)
$$E\left[\begin{array}{c|c} g(j,t) & y(j,t) \\ g(j,t-1) & y(j,t-1)\end{array}\right] = \begin{pmatrix} \iota' & \rho\cdot\iota' \\ \rho\cdot\iota' & \iota'\end{pmatrix}\Omega^{-1}B^{-1}\begin{pmatrix} y(j,t)-\alpha(t) \\ y(j,t-1)-\alpha(t-1)\end{pmatrix}$$
$$V\left[\begin{array}{c|c} g(j,t) & y(j,t) \\ g(j,t-1) & y(j,t-1)\end{array}\right] = \begin{pmatrix} 1 & \rho \\ \rho & 1\end{pmatrix} - \begin{pmatrix} \iota' & \rho\cdot\iota' \\ \rho\cdot\iota & \iota'\end{pmatrix}\Omega^{-1}\begin{pmatrix} \iota' & \rho\cdot\iota' \\ \rho\cdot\iota & \iota'\end{pmatrix}$$

where B is a block-diagonal matrix with B(t) and B(t-1) on the diagonal, and ι is a Kx1 vector of ones. The covariance matrix Ω has the following block form:

$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$, with $\Omega_{11} = \iota\iota' + \Sigma(t)$, $\Omega_{12} = \Omega_{21}' = \rho\iota\iota' + R\Sigma(t)^{1/2}\Sigma(t-1)^{1/2}$, and $\Omega_{22} = \iota\iota' + \Sigma(t-1)$.[15]

The conditional mean and variance in Equation (5) are just the two-period generalizations of the estimates of governance and their precision based on the one-period unobserved components model that we used in Section 2, i.e. Equation (5) is the exact analog of Equations D2 and D3 in Appendix D. In fact, if we set $\rho = r_k = 0$ for all sources k, then we recover exactly the estimates of governance that we had before. The advantage of this two-period formulation is that we now have specified the joint distribution of governance in the two periods for each country, conditional on the observed data in the two periods. Since we have modeled the joint distribution over the two periods of governance, we can base inferences about governance in the two periods, as well as changes in governance, on this joint distribution. We also note that the discussion of inference about changes over time in governance based on individual indicators in the previous section is just a special case of this more general formulation.[16]

We implement this two-period model using our actual dataset, over the period 1996-2004. We restrict attention to a balanced set of sources that are available in both periods for the two indicators. In order to implement this calculation, we need to have estimates of the parameters of the model in both periods (the α's, β's, and σ's), as well as estimates of the correlation over time of the errors in the individual sources (the r's) and the correlation of unobserved governance itself, ρ. We obtain these parameters in

---

[15] To obtain Equation (5), note that the (2K+2)x1 vector $(g(t), g(t-1), y(t), y(t-1))'$ is normally distributed with mean $(0, 0, \alpha(t), \alpha(t-1))'$ and variance-covariance matrix V with the following block form: $V_{11} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $V_{12} = \begin{pmatrix} \iota' & \rho \cdot \iota' \\ \rho \cdot \iota' & \iota' \end{pmatrix} B$, and $V_{22} = B\Omega B'$. Standard results for the partitioned multivariate normal distribution imply that the distribution of governance conditional on the observed data is normal with mean and variance given by Equation (5).

[16] To see this, set the number of sources K=1 and assume that $\alpha(t)=0$, $\beta(t)=1$, and $\sigma(t)=\sigma$ for this one source. Equation (5) then gives the conditional mean and variance of the level of governance in the two periods based on this single source. The expected change in governance conditional on the data is then just the difference between the conditional means in the two periods, and the conditional variance of the change is just the sum of the variances in the two periods less twice the covariance.

two steps. First, we estimate the one-period unobserved components model in 1996 and in 2004, to obtain estimates of the α's, β's, and σ's. We refer to this as the "static model" estimates. We also retrieve the estimates of governance and standard errors from the static model, to use as a basis for comparisons with the two-period model. Second, we calculate the correlation over time of these static estimates of governance as an estimate of ρ. In this second step we also insert the estimated parameters of the static model into Equation (4) and retrieve estimates of the errors in the sources in the two periods as residuals. The correlation over time in these estimated residuals serves as our estimate of the correlation in the errors. We then insert all the estimated parameters, together with the data, into Equation (5) to obtain our final estimates of governance in the two periods conditional on the data, as well as the variance-covariance matrix of these estimates. We refer to these as the "dynamic model" estimates.

Table 7 summarizes the results of this calculation for the six governance indicators. In the top panel we present some summary statistics to aid in the comparison of governance estimates based on the single-period, or static model, and the two-period, or dynamic model. In the first two columns we report the correlation between the estimates of governance based on the static and dynamic models, in the two periods, 2004 and 1996. These correlations are virtually one for all six indicators in both periods, suggesting that our estimates of the levels governance do not change very much if we take into account persistence in governance and in the error terms. The third column reports the correlation of the change over time in the estimates of governance according to the two models. In light of the high correlations in levels between the two models, it is not very surprising that the correlation of changes is also very high, averaging 0.93 across the six indicators.

The next two columns of Table 7 report the average absolute change in the governance estimates for the static and dynamic models. These changes are roughly half as large in the dynamic model than in the static model, averaging 0.17 and 0.32 respectively. The reason the dynamic model gives much smaller estimates of the change in governance over time is because the estimated persistence in governance is quite strong relative to the estimated persistence in the error terms. Averaging across the six indicators, the persistence in unobserved governance is estimated to be 0.89.

This is over twice as large as the persistence in the error terms, which averages 0.42 across all sources and indicators. Based on our intuitions from the simple example above, we should expect to find substantially smaller estimates of the change in governance when we take this pattern of persistence into account, and this is in fact what happens.

The bottom panel of Table 7 summarizes the consequences of this persistence for inference about changes in governance. Formally our objective is to test the null hypothesis that the change in unobserved governance is zero conditional on the observed data. We begin by calculating the z-statistic associated with this hypothesis for each country, using the static and dynamic models. For the static model, we simply take the absolute change in our estimate of governance, and divide by the square root of the sum of the variances of the estimate of governance in the two periods. For the dynamic model, we calculate the variance of the change in governance as the sum of the estimated variances in the two periods, minus twice the estimated covariance between the two periods. The square root of this variance becomes the denominator of the z-statistic for the dynamic model. The average z-statistics are smaller in the dynamic model than in the static model, again consistent with the intuitions developed above. For the static model, the z-statistics average 0.82, as opposed to 0.59 for the dynamic model. This in turn implies fewer statistically significant changes in governance based on the dynamic model, as reported in the next two columns. The average number of significant changes at the 10 percent level falls by half from 21 to 10 once we take persistence into account.

Although a relatively small number of changes in the aggregate indicators signal statistically significant changes in unobserved governance, it is worth noting that the proportion of significant changes is much higher for the aggregate indicator than it is for individual indicators. Recall from the previous subsection that only the top one percent of changes in an individual indicator with typical persistence in unobserved governance and the error term would be significant at the 90 percent level. This is not because individual indicators do not register large changes for individual countries – in fact frequently they do so. Rather, it is because the margins of error associated with changes in individual data sources are large. In contrast, for the aggregate indicators we find that between five and seven percent of all changes signal statistically significant

changes in governance at the same significance level, reflecting the greater precision of the aggregate indicators. This illustrates the benefits of aggregation for assessing changes over time, as well as levels, of governance.

We also note that a substantially larger proportion of changes in governance are significant if we relax the standard of significance to 75 percent, for example. For the case of a typical individual indicator, we have already seen that the top seven percent of changes in the observed data would signal significant changes in unobserved governance. For our composite indicators this fraction is higher. For example, for the Voice and Accountability measure, seven percent of the changes are significant at the 90 percent level, while 12 percent of changes, or 23 cases, would be significant at the 75 percent level. Finally, these calculations somewhat understate the number of significant changes because they are based on a subset of our data sources that are available in both periods – had more of our sources in 2004 been available in 1996, we would have had even more significant changes over time.[17]

Finally, it is useful to compare the statistically significant changes in governance identified by the dynamic model with the "large" changes in governance we identified in Section 2.3 of this paper using a very simple rule of thumb. We begin by identifying all changes in governance based on the static model for which the 90 percent confidence intervals in the two periods do not overlap, as per the rule of thumb. Note that this is a more stringent condition for identifying significant changes in governance than the t-tests for the static model we have just discussed.[18] On average, there are nine significant changes in governance per indicator according to this rule of thumb applied to the simple static model, as compared with 10 in the dynamic model. There is a remarkable degree of overlap between the significant changes identified by the rule of thumb and the dynamic model. On average, eight of the nine changes identified by the rule of thumb are also significant in the dynamic model. Moreover, comparing the second and third-

[17] We have also analyzed changes over the period 1998-2004, and find a similar proportion of changes to be statistically significant. While on the one hand we are looking at changes in governance over a shorter period of time, on the other hand we have more data sources available in both periods on which to base our assessment of changes.

[18] Requiring 90 percent confidence intervals not to overlap is equivalent to requiring the absolute change in estimated governance to be larger than the sum of the standard errors in the two periods. This sum is always larger than the square root of the sum of the squares of these standard errors.

last columns of this panel, it is clear that the dynamic model turns up very few significant changes not identified by the rule of thumb. Although the simple rule of thumb and the more formal model turn up more or less the same set of significant changes in governance, it is important to note that the magnitude of these changes is substantially smaller in the formal dynamic model.

In summary, we have developed a dynamic version of the single-period unobserved components model that we have used to construct our aggregate governance indicators. The advantage of specifying a dynamic version of the model is that it allows us to make formal statistical inferences about changes in unobserved governance based on our changes in the composite governance indicators. But this advantage comes at a cost. The two-period model is substantially more complicated to implement, particularly when the set of underlying data sources is not the same in both periods. Given that the number of data sources we use has expanded substantially over time, this is a significant limitation. Fortunately, however, we have seen that using a simple rule of thumb for identifying large changes over time in our static or single-period estimates of governance corresponds quite closely to formal inference regarding the significance of changes in governance. Because of this, we continue to use the single-period unobserved components model to construct the aggregate governance indicators in each period, and recommend using the simple rule of thumb that 90 percent confidence intervals do not overlap for identifying changes in governance that are likely to be statistically significant.

4. De Jure vs. De Facto Measures of Governance

A recurrent theme in this paper is that individual sources of governance data are imperfect and provide only noisy signals of unobserved governance. We emphasize at the outset that this problem is not unique to the subjective or perceptions-based measures of governance on which we rely. Rather, it is pervasive in all efforts to measure governance, or any other socioeconomic variable for that matter. What are the sources of this measurement error? In the case of our governance data, we emphasize two distinct sources. First, as always, specific concepts may be imperfectly measured. Survey responses to a question such as "is it difficult to start a business?" reflect sampling variation in the survey. Expert assessments of the difficulty of starting a business rely on the imperfect information available to such experts, and hence also contain measurement error. Second, and perhaps more important, is that there are inevitably gaps between the specific concept being measured and the broader notion of governance that it is intended to proxy. For example, the ease of starting a business is just one of many dimensions of the regulatory environment, and as such is an imperfect proxy even if the narrow concept of business entry regulation were perfectly measured.

This broad notion of measurement error clearly also applies to "objective" or quantifiable measures of governance. Consider for example the very useful "Doing Business" project of the World Bank, which has compiled objective measures of various dimensions of the regulatory environment across countries, by interviewing law firms around the world about formal rules and regulations in their countries. These measures are subject to the same two sources of measurement error. As always there may be gaps between the *de jure* rules on the books, and their *de facto* application. And as with the subjective measures, there are gaps between this specific dimension of regulation and the overall quality of the regulatory environment. The same limitations apply to many other objective measures of governance that have been proposed. Trade taxes as a share of total tax revenue has been suggested as a proxy variable for the ability and willingness of the government to broaden its tax base. This measure is also subject to measurement error given the dubious quality of data on public finances in many developing countries, and moreover is an imperfect proxy of a government's fiscal capability. Similarly, although it is easy to observe whether a country has an independent anti-corruption commission, it is much more difficult to measure whether such a commission is in fact independent or empowered to act.

Although objective indicators of governance are subject to measurement error, this uncertainty is rarely quantified or made explicit. In an earlier paper we made an effort to quantify the margins of error associated with several leading objective indicators of governance.[19] We found that this broad notion of measurement error was as important for objective indicators as for the subjective indicators we develop. We did not however attempt to distinguish between the two sources of measurement error: difficulty in measuring specific concepts, on the one hand, and the gap between specific concepts and broader notions of governance, on the other. In this section of the paper we make an effort to focus on the first source of measurement error. In particular, we focus on understanding the gaps between *de facto* perceptions of quite specific dimensions of governance, and the corresponding *de jure* regulations.

We consider two measures of the *de facto* environment facing firms, taken from the survey of over 8000 firms in 104 countries carried out by the World Economic Forum in 2004 as an input to their Global Competitive Report. These two variables capture firms' assessment of the ease of starting a business, as well as their reported tax burden.[20] We match these with two closely-related *de jure* measures from other sources. For ease of starting a business, we draw on the Doing Business project at the World Bank discussed above. From this dataset we take the number of days required to start a business. For perceptions of the tax burden, we have independently collected statutory tax rates for the sampled countries, and within it, for the types of firms by sector, and mapped these rates into the firm level data. We then aggregate these up to the country level to obtain average measures of the statutory tax burden.[21]

[19] See Kaufmann, Kraay and Mastruzzi (2004).
[20] For the past number of years, collaboration between WBI and the WEF has resulted in an in-depth coverage of governance in the survey, and in the WBI contribution of a governance chapter for each GCR. For details on the data we use for the text described above, and the related coverage of these governance issues at the micro-level, see the Governance chapter in the GCR 2004, at http://www.worldbank.org/wbi/governance/pubs/gcr2004.html.
[21] The main source for the effective tax rates was the PricewaterhouseCoopers report "Corporate taxes: worldwide summaries (2003-2004)", covering 85 of our sample of 104 countries. As some countries have differential tax rates, to map the country-level data from the report to the individual firm-level data from the GCS we used, in addition to country criterion, individual characteristics such as size, sector, and whether the firm exports or not. For those countries for which the report has no information we used the country average calculated by KPMG in their "Corporate tax rate survey".

We begin with simple ordinary least squares regressions of perceptions of ease of starting a business on the corresponding objective measure (first column of Table 8). Not surprisingly, the objective measure enters negatively and is highly statistically significant with a t-statistic of more than five, indicating that firms perceive it more difficult to start a business in countries where the number of days required to do so is large. More interesting for our purposes is the observation that the R-squared of the regression is very modest, at only 0.23.

We cannot say at this point whether this reflects measurement error in the subjective or the objective measure, as either one would contribute to a low R-squared. One hypothesis however is that the objective measure fails to capture the extent to which the formal requirements to start a business are altered by the presence of corruption or other forms of informality in their application. To investigate this possibility we add our aggregate measure of Control of Corruption to the regression.[22] We find that this variable enters positively and highly significantly, indicating that perceptions of the ease of starting a business are significantly better in countries with less corruption, even after controlling for the *de jure* rules governing business entry. Once we add corruption, the coefficient on the *de jure* rules falls by half, and its significance also drops to the 10 percent level. Moreover the adjusted R-squared of the regression doubles to 0.44, indicating substantial explanatory power for this additional variable.

There is however an obvious difficulty with this result. It could well be the case that firms' responses to the question regarding business entry are non-specific, in the sense that they will provide low responses if their assessment of the overall business environment is negative. This generalized dissatisfaction could account for the significance of the corruption variable, rather than the extent to which business entry procedures are tainted by corruption. We address this possibility in the next three columns. One test for this problem of non-specificity is to ask whether unrelated objective measures of the business environment also predict perceptions about ease of entry. We do this in the third column by adding the objective tax burden question to the regression. If firm responses reflect generalized dissatisfaction, we might expect this variable also to enter significantly, yet it does not. In the fourth column we instead add

[22] Recognizing that the dependent variable is one of many individual data sources entering in the regression, we lag the corruption measure and use the 2002 version.

firms' responses to a question about the overall regulatory environment that they face. Again we find that corruption remains highly significant, and in this case the general question about regulation is also highly significant. This suggests that while non-specificity of responses may be a concern, it does not fully account for the significance of the corruption measure in the previous specifications. Interestingly, in both specifications, we find that the coefficient on the objective entry measure becomes larger and more significant as we add these control variables. Finally we note that all these results go through when we put all four variables in the regression.

The second and third panels of Table 8 reveal interesting differences between developing countries on the one hand, and OECD and newly-industrialized countries, on the other. In the developing country sample, the results described above go through for the most part. However, it is interesting to note that the magnitude and significance of the objective measure is in general smaller in the developing country sample, and larger in the industrial country sample, while the converse is true for the corruption variable. Taken together these results suggest that firm perceptions of the ease of starting a business depend on both *de jure* rules, as well as the institutional environment in which those rules are applied. Moreover, the relative importance of *de jure* rules seems to be higher in industrial than in developing countries. More broadly, the lesson from this simple exercise is that it can be misleading to rely exclusively on either perceptions of *de facto* governance or objective measures of the *de jure* rules.

We perform the same sequence of regressions using the question on perceptions of tax burdens from GCS as the dependent variable. The results are broadly similar to those discussed above, and are reported in the continuation of Table 8. In the full sample of countries, we find that perceptions of tax burdens are strongly correlated with our *de jure* measure of statutory tax rates. While in the full sample of countries we do not find corruption to enter significantly, it does in the developing country sample where we might expect corruption to matter more for perceptions of the tax burden. As before, we address the possibility that the tax burden question captures generalized dissatisfaction rather than a specific concern with taxation by including the objective measure of days to start a business, and we find that the corruption variable remains significant. Also consistent with our priors, we find that differences in statutory tax rates have much stronger explanatory power for perceptions of tax burden in the

industrial country sample. While the overall results are not quite as strong as for the business entry example discussed above, qualitatively the picture that emerges is quite similar.

In sum, the results suggest that assessments of governance should not be based solely on objective measures of the *de jure* situation. We have seen that firms' perceptions of the ease of starting a business, and the weight of their tax burden, depend not only on the *de jure* regulations that they face, but also on the environment in which these regulations are applied. Many laws and regulations are often adopted, yet implementation is subverted due to the many informal mechanisms that often prevail. In these settings frequently the essence of how policies and regulations are actually implemented may be missed by objective indicators. This is not to say, of course, that firm-based surveys of perceptions are devoid of margins of error and related challenges. Rather, the results we have shown emphasize the importance of relying on a range of measures to assess governance, and on recognizing that no single measure is a perfect proxy for governance.

5. Interpreting Governance-Income Correlations

In this section of the paper we briefly discuss two methodological issues that arise in interpreting the strong positive correlation observed between measures of governance and per capita income across countries. We first consider – and discount – the possibility that these strong correlations are a consequence of "halo effects", i.e. an upward bias in perceptions of governance in rich countries simply because they are rich. We also discuss – and refute – the argument that the weak governance performance of countries in Africa should be discounted in some sense because these countries are poor.

5.1 Halo Effects

Perceptions-based measures of governance such as the ones we develop here are potentially subject to a number of biases. One common critique is that perceptions of governance are biased upwards in rich countries because respondents view the development success of the country in question as evidence that institutional quality is good. This type of bias is sometimes referred to as a "halo effect".[23] This in turn implies that part of the observed high correlation between per capita incomes and governance spuriously reflects this bias.

To formalize the idea of halo effects, suppose that we can write our observed estimates of governance, g*, as the sum of true governance, g, and an error term, u:

(6) $$g^* = g + u$$

The essence of the halo effect argument is that this error term u is correlated with per capita incomes, y. The relevant question then is the extent to which this spurious correlation can account for the high observed correlation between measured governance and per capita incomes. Intuitively, it should be clear that in order for halo effects to substantially account for the correlation between incomes and measured governance, it

[23] A recent statement of this critique can be found in Glaeser , La Porta, Lopez-de-Silanes, and Shleifer (2004), who assert that much of the correlation between subjective measures of governance and levels of development is attributable to this type of bias.

must be the case that the correlation between the error and income is large. Perhaps less obviously, it must also be the case that the variance of the error term is large relative to the variation in governance. Otherwise, even if the error term is strongly correlated with income, the fact that it accounts for little of the variance in measured governance means that it will have little impact on the correlation between measured governance and per capita income. Our argument in a nutshell is that for reasonable assumptions on the importance of measurement error, this measurement error would have to be implausibly highly correlated with per capita incomes in order to constitute a significant source of bias.

To formalize this intuition, we decompose the observed correlation between measured governance and per capita income into a term reflecting the true correlation between governance and income, and a term attributable to the halo effect:

$$(7) \qquad \mathrm{CORR}(g^*, y) = \sqrt{1-s} \cdot \mathrm{CORR}(g, y) + \sqrt{s} \cdot \mathrm{CORR}(u, y)$$

where $s = V[u]/V[g^*]$ is a measure of how noisy the governance indicator is. Note also that the correlation between measured governance and per capita income that we see in the data is around 0.8.

To understand this expression, suppose that the true correlation between governance and income were zero, so that all of the observed correlation between income and governance is due to the second term capturing halo effects. This consists of two ingredients: the actual correlation of the error term with per capita income, which is multiplied by the square root of the share of the variance in governance due to the error term. Suppose that the governance indicator is very noisy so that the share of the variance approaches one. Then the correlation of the error term with per capita income must be equal to the observed correlation in the data. Suppose however that the governance indicator is at least somewhat informative, so that s is less than one. Then in order to match the observed correlation in the data, the halo effect correlation in the error term must be even larger than the 0.8 observed in the data. This example illustrates how the importance of halo effects in accounting for the observed correlation between governance and per capita income depends on both the strength of the halo

effect itself, as well as the relative importance of measurement error in the governance indicator.

This example is extreme because we have assumed that the true correlation between governance and income is zero. We now relax this assumption and revisit the question of how strong halo effects need to be to account for the observed correlation between measured governance and per capita income of 0.8. We do this with the help of Figure 5, which graphs the strength of the halo effect, i.e. CORR[u,y], on the vertical axis, against the share of the variance in governance due to the residual, i.e. s, on the horizontal axis. The different lines on the graph correspond to different assumptions for the true correlation between governance and income. We have already discussed the intuition for the case where this correlation is zero, shown as the highest line in the graph. When the share of the variance in governance due to measurement error is one, the halo effect correlation must be equal to 0.8. As we move to the left and the governance indicator becomes more informative, the required correlation increases.

The lines corresponding to successively higher true correlations between governance and income fall everywhere below the first series. This is because once we allow for some correlation between true governance and income, the halo effect needed to account for the correlation between observed governance and income is weaker. Interestingly, however, even if the true correlation is quite substantial at 0.6, the lowest line in Figure 5 tells us that halo effects must still be quite considerable, with a correlation of at least 0.5, to match the observed data.[24] This lower bound occurs for intermediate values of the share of the variance of governance due to measurement error. It is also interesting to ask what a reasonable value for this share might be, in order to pin down more precisely how strong halo effects must be. One way to do so is to consider the standard errors of the governance estimates, which average around 0.25 as compared with the standard deviation of measured governance of 1. This suggests that the share of the variance of governance due to the error term is in fact quite small at $s=0.25^2=0.06$. For this low variance share, the halo effect correlation would need to be

[24] We do not consider higher values for the true correlation than 0.6. This is because we are trying to see the extent to which halo effects might result in an observed correlation of 0.8 which is substantially higher than the true correlation. If the observed correlation and the true correlation are close to each other, then the halo effects argument becomes unimportant empirically.

0.9 in order to match the observed data. If the true correlation between governance and income were much lower, for example at 0.4, then even if measurement error in governance were perfectly correlated with per capita income it would not be possible to generate the observed correlation between governance and per capita incomes.

This strong conclusion is driven by the assumption that that measurement error accounts for a relatively small portion of the variation in observed governance. As a result this measurement error needs to be very highly correlated with incomes in order to match the data. One could argue that we are understating the importance of measurement error by relying on the estimated standard errors from our governance indicators. After all, these are based on the assumption that measurement error is uncorrelated across different sources of governance data. However, if halo effects are important, the measurement error in individual sources will be correlated not only with per capita income, but also with each other. This in turn would imply a greater imprecision of the governance estimates. To capture this possibility, suppose that the standard error of the governance estimates were twice as large as what we actually have, at 0.5.[25] This implies s=0.25, and for this value of s we can see from Figure 5 that the halo effect correlation would still need to be very high at almost 0.6 in order to match the data.

In summary, these results suggest to us that although halo effects may well be present in perceptions-based measures of governance, these halo effects need to be implausibly strong in order to impart a substantial upward bias in the correlation between measured governance and per capita incomes. Moreover, it is worth noting that there may well be other factors offsetting such halo effects. One is the tendency of survey respondents in developed countries to be particularly critical of their own institutions.[26] It is also worth noting that some cross-country polls of experts deliberately apply higher standards to rich countries when assessing their governance.[27] Overall, then, we do not

[25] In Kaufmann, Kraay and Zoido-Lobaton (1999), Table 5, we show that the estimated margins of error would be roughly twice as large if we assume that the correlation of error terms across sources is 0.5 instead of 0.

[26] For treatments of these effects in survey data, see Kaufmann and Wei (1999) and Hellman, Kaufmann and Schankerman (2000)

[27] For example, in our discussions with PRS, we learned that this source penalizes rich countries that in their view have the resources to reduce corruption but fail to do so.

think that halo effects are a significant source of bias in the correlations between governance and per capita incomes our data.[28]

5.2 Controlling for Income in Governance Comparisons

In a recent paper, Sachs and others (2004) have argued that weak governance is not a major factor in Africa's poor growth performance. The argument is that, once we control for per capita income, countries in sub-Saharan Africa do not have particularly poor governance indicators. This point is illustrated in Figure 6, which plots our 2004 Rule of Law measure (on the vertical axis) against the logarithm of real per capita GDP in the mid-1990s (on the horizontal axis). Note that the per capita income variable has been rescaled to have mean zero and standard deviation of one, as does the governance indicator. Countries in Sub-Saharan Africa are highlighted in red. A striking observation from this graph is that over half (27 out of 46) of the countries in the region actually fall above the simple ordinary least squares regression line, shown in black. At first glance, this appears to lend credence to the argument that governance in Africa is on average what one might expect given the region's low income levels.

However, it is misleading to conclude from this simple graph that Africa's governance performance is reasonable given its per capita income. This interpretation of the graph is valid only to the extent that the OLS regression line would capture a causal relationship from higher income to better governance. But a large body of research indicates that there is substantial causation in the other direction as well – better governance leads to higher incomes. Moreover, the magnitude of the estimated effect of governance on per capita incomes in the long run is large.[29] Available estimates suggest that a one standard deviation improvement in governance would lead to a two- to three-fold difference in income levels in the long run. A one standard deviation change in governance would correspond to, for example the difference between Kenya

[28] It is of course possible that halo effects are associated with countries' recent growth performance, rather than with income levels. We can use the analysis of this section to consider this case as well. The main insight is that since the correlation between recent growth and governance is typically fairly modest, growth-related halo effects would not need to be as large in order to impart a proportionately larger bias to this correlation.

[29] See for example Hall and Jones (1999), Acemoglu, Johnson and Robinson (2000), Kaufmann and Kraay (2002), Alcala and Ciccone (2004), Rodrik, Subramanian and Trebbi (2004), and Rigobon and Rodrik (2004).

and Turkey on our 2004 Rule of Law indicator. This means that the simple OLS relationship will exaggerate the positive effects of income on governance because it also reflects the strong effect in the opposite direction, from governance to incomes. In order to compare governance in Sub-Saharan Africa to what might be expected given income levels, we need to first isolate these two directions of causation, so to be able to focus in particular on the causal effect of income on governance.

The red and green lines in Figure 6 show two alternative estimates of the causal effect of income on governance. The (slightly) upward-sloping one, in red, comes from Rigobon and Rodrik (2004). They study the causal relationships between per capita income, democracy, rule of law, openness to international trade, and geography, using identification through heteroskedasticity to isolate the causal effects.[30] As expected, this red line is substantially flatter than the OLS regression line, consistent with the intuition that the latter relationship overstated the true causal effect of incomes on governance. This flattening has important consequences for our conclusions about the quality of governance in Africa controlling for income levels. Once we isolate this much weaker effect of income on governance, we find that only 7 out of 46 countries in the region fall above the regression line: Ghana, Lesotho, Cape Verde, Namibia, South Africa, Botswana, and Mauritius. In contrast, the vast majority of countries in Africa have governance that is worse than their income levels would predict.

The weakly downward-sloping green line presents another estimate of the effect of income on governance, coming from Kaufmann and Kraay (2002). In this paper we used a different approach to identification and found a zero or even negative impact of income on governance. While this finding may be somewhat extreme, it leads to the same conclusions regarding the quality of governance in Africa – now only 6 out of 46 countries in the region fall above the regression line, indicating governance levels better than what per capita incomes would predict.

[30] We use their specification excluding democracy, which implies that a one standard deviation increase in log per capita GDP improves rule of law by 0.14 standard deviations. They use a different measure of rule of law for the mid-1990s taken from Knack and Keefer (1995). However, its correlation with our rule of law indicator is above 0.8, so we can reasonably use the estimated coefficient from this paper with our governance indicator, suitably standardized. Note also that in the system of equations estimated by Rigobon and Rodrik (2004) the conditional expectation of governance given per capita income also reflects the indirect effects of income on openness, which in turn affects the rule of law. However, these estimated indirect effects are so small that our conclusions are essentially unaffected by ignoring them.

Overall this evidence suggests that it would be inappropriate to discount the governance performance of countries in Sub-Saharan Africa based on their low income levels. The reason is simple. The only way to justify such a discount is to argue that higher incomes exert a positive causal effect on governance. But available evidence suggests that the causal impact of incomes on governance is small. Rather, the observed correlation between governance and per capita incomes primarily reflects causation in the other direction: better governance raises per capita incomes.

5. Conclusions

There is by now broad consensus among academics and policymakers alike that good governance matters for economic development. There is also growing awareness in the aid community that good governance matters for the effectiveness of development assistance. In light of this it is important to be able to measure levels and changes over time in governance across countries. This paper represents the latest installment of our aggregate governance indicators which seek to provide such information. Relative to previous years, these indicators reflect a significant expansion of our underlying data set of several hundred individual variables measuring perceptions of governance, drawn from 37 separate data sources.

In our work we have emphasized the difficulty of measuring governance. We have argued that one of the strengths of our composite governance indicators is that they can be more informative than individual data sources: on average the aggregation reduces the margin of error by about one-half. Further, given the increasing number of separate data sources now at our disposal to construct these aggregate indicators, we find that the margins of error of the latest period under measure are smaller than in earlier periods. However, these margins of error, even in our most recent aggregate indicators, still remain substantial, and thus all our previous cautionary suggestions regarding interpretation continue to apply.

At the same time, we have emphasized that these margins of error are not unique to perceptions-based measures of governance, but are an important feature of all efforts to measure governance. In fact, in previous work we have documented that objective measures also have substantial margins of error. Moreover, we believe that the type of perceptions data on which we rely provides insights into governance that are difficult to obtain from more objective or quantifiable measures. For example, we show that firm's perceptions of the difficulty of starting a new business, or of their tax burdens, do not depend solely on the relevant legal framework governing business entry and taxation. Rather, firms views on these issues are also importantly influenced by the degree of corruption in their country (particularly so in developing countries), suggesting that not only do formal rules matter, but also the institutional environment in which these rules are applied and enforced. Thus, wherever objective data on governance or

investment climate is collected (such as de jure data of the numbers of steps required by the regulations to start a business), a comprehensive analysis of governance and institutional change ought to be complemented by data from the reports of the economic agents on the ground, such as firms or users of services, which inevitably will contain an element of subjectivity.

Policymakers are often particularly interested in *trends* in institutional quality: is governance improving or worsening over time in a particular country? As we have emphasized in our work, the presence of measurement error in all types of governance indicators, including our own, makes assessing trends in governance a challenging undertaking. In this paper we developed a formal statistical methodology, as well as a simple rule of thumb, for identifying changes in governance that are likely to be statistically and practically significant. Over the eight-year period from 1996-2004 spanned by our governance indicators, we find that in about five to seven percent of countries we can be confident (at the 90 percent significance level) that governance has changed substantially. And at a lower 75 percent significance level, roughly 20 percent of all observed changes stand out as significant. Importantly, we show that there is a great deal of agreement among our many data sources about the direction of change in governance in these countries. Overall this reminds us that while often institutional quality changes takes place haltingly, gradually, or not at all, there are also countries where one can point to sharp improvements or deteriorations even over a fairly short eight-year period. Significant and rapid institutional change, while not the norm, is feasible and does take place in practice.

Finally, we have discussed two important issues that arise in interpreting the strong positive correlation between governance and income levels. Some observers have argued that these positive correlations are substantially due to “halo effects” – perceptions of governance in rich countries are good simply because the countries are rich. We have argued that such halo effects would need to be implausibly large to account for cross-country correlations between governance and incomes.

We have also considered the frequently-heard argument that poor levels of governance be significantly discounted where the country is poor. Put differently, to what extent does it make sense to ask whether a country is well or poorly governed

given its income level? This issue is often raised in the context of Sub-Saharan Africa, where too many countries are both very poor, and very poorly governed. We make the simple observation that in order to answer this question, it is necessary to isolate the causal impact of income levels on governance. Simply relying on the observed correlation is inappropriate, as much of this reflects strong causal effects running from governance to per capita incomes. While identifying the effects of income on governance is difficult, the few available estimates suggest that this feedback effect is minimal. As a result, there is little basis on which to argue that the poor governance performance many countries in sub-Saharan Africa should be discounted simply based on low income levels.

In conclusion, it is important to keep some perspective on this contribution. While these aggregate governance indicators have been useful in providing a general snapshot of the countries of the world for various broad components of governance, now for 8 years, and while the margins of error have declined over time, they remain a rather blunt instrument for specific policy advice at the country level. As we have argued in the past, these aggregate indicators need to be complemented with in-depth in-country governance diagnostics, based on micro-surveys of households, firms and public officials within the country. The lessons being drawn from these combined aggregate and micro-data sets do point to the importance of moving concretely to the next stage of governance reforms, in Africa and elsewhere. These, among others, are to stress reforms in transparency (such as natural resource revenue transparency mechanisms, disclosure of assets of politicians, voting records of parliamentarians, political campaign contributions, and fiscal accounts), in altering incentives in institutions so to increasingly focus on prevention and deterrence (rather than overly relying on prosecutions), and in working more closely with other key actors outside the public sector as well, such as the heretofore neglected private sector.

References

Acemoglu, Daron, Simon Johnson, and James Robinson. "The Colonial Origins of Comparative Development. *American Economic Review*. 91(5):1369-1401.

Alcala, Francisco and Antonio Ciccone (2004). "Trade and Productivity". *Quarterly Journal of Economics.* 119(2):613-46.

Efron, Bradley and Carl Morris (1971). "Limiting the Risk of Bayes and Empirical Bayes Estimators – Part 1: The Bayes Case." *Journal of the American Statistical Association* 66:807-15.

——— (1972). "Limiting the Risk of Bayes and Empirical Bayes Estimators – Part 2: The Empirical Bayes Case." *Journal of the American Statistical Association* 67:130-39.

Goldberger, Arthur (1972). "Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables." *International Economic Review* 13(1):1-15.

Glaeser, Edward, Rafael LaPorta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2004). "Do Institutions Cause Growth?". Journal of Economic Growth.

Hall, Robert E., and Charles Jones (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114(1):83-116.

Hellman, Joel, Geraint Jones, Daniel Kaufmann and Mark Schankerman (2000) 'Measuring Governance, Corruption, and State Capture: How Firms and Bureaucrats Shape the Business Environment in Transition Economies', World Bank Policy Research Working Paper No. 2312, Wash. DC.

Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobatón (1999a). "Aggregating Governance Indicators." World Bank Policy Research Working Paper No. 2195, Washington, D.C. http://www.worldbank.org/wbi/governance/pubs/aggindicators.html.

——— (1999b). "Governance Matters." World Bank Policy Research Working Paper No. 2196, Washington, D.C. http://www.worldbank.org/wbi/governance/pubs/govmatters.html.

——— (2002). "Governance Matters II – Updated Indicators for 2000/01." World Bank Policy Research Working Paper No. 2772, Washington, D.C. http://www.worldbank.org/wbi/governance/pubs/govmatters3.html.

Kaufmann, Daniel and Aart Kraay (2002). "Growth Without Governance". Economia. 3(1):169-215.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2004). "Governance Matters III: Governance Indicators for 1996, 1998, 2000, and 2002". World Bank Economic Review. 18:253-287.

Kaufmann, Daniel, and Shang-Jin Wei, 'Does Grease Money Speed Up the Wheels of Commerce?' World Bank Policy Research Working Paper No. 2254, Wash. DC, December 1999

Knack, Steven and Philip Keefer (1995). "Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures." *Economics and Politics*, 7, 207-227.

Rigobon, Roberto and Dani Rodrik (2004). "Rule of Law, Democracy, Openness, and Income: Estimating the Interrelationships. Manuscript. MIT and Kennedy School

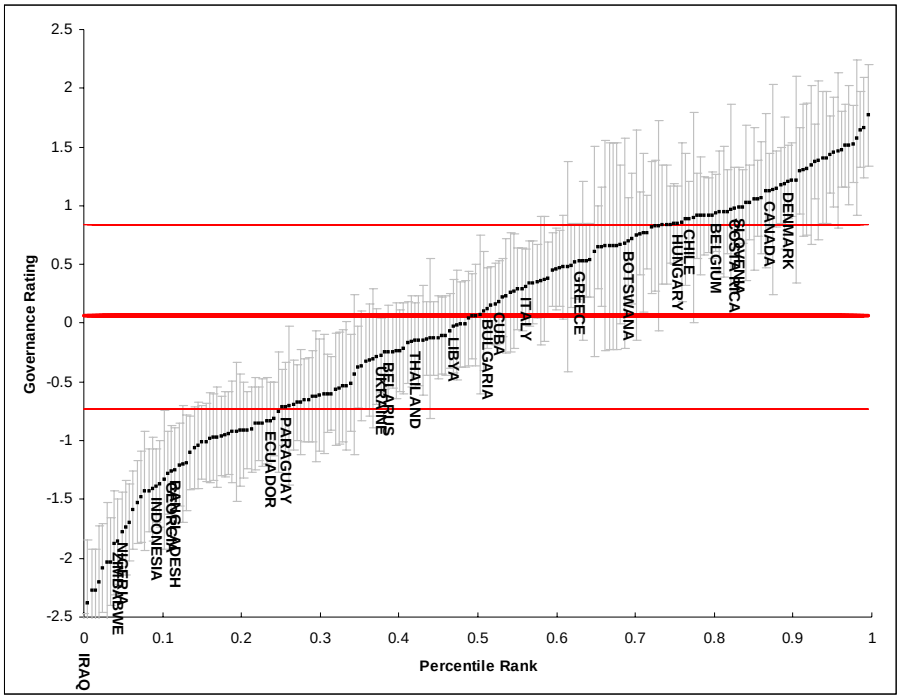
Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi (2004). "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development". *Journal of Economic Growth* 9(2):131-165.

Sachs, Jeffrey, John W. McArthur, Guido Schmidt-Traub, Margaret Kruk, Chandrika Bahadur, Michael Faye, and Gordon McCord, (2004), "Ending Africa's Poverty Trap", Brookings Papers on Economic Activity, 2004(1).

World Bank (2004). "Doing Business in 2004: Understanding Regulation". Washington: Oxford University Press.

Figure 1: Margins of Error for Governance Indicators, 2004

Political Stability and Absence of Violence



Control of Corruption

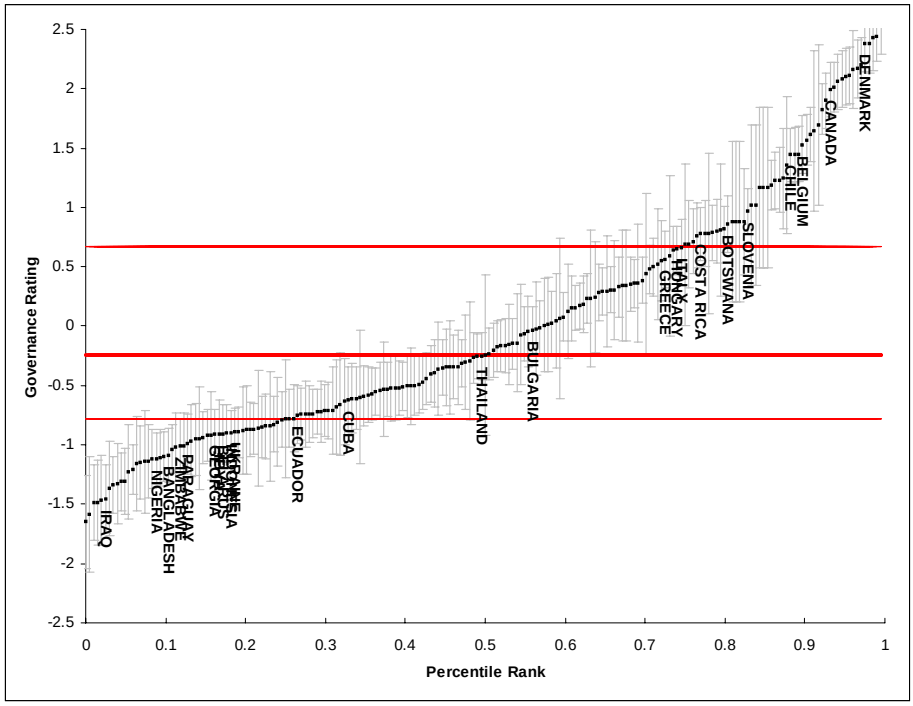
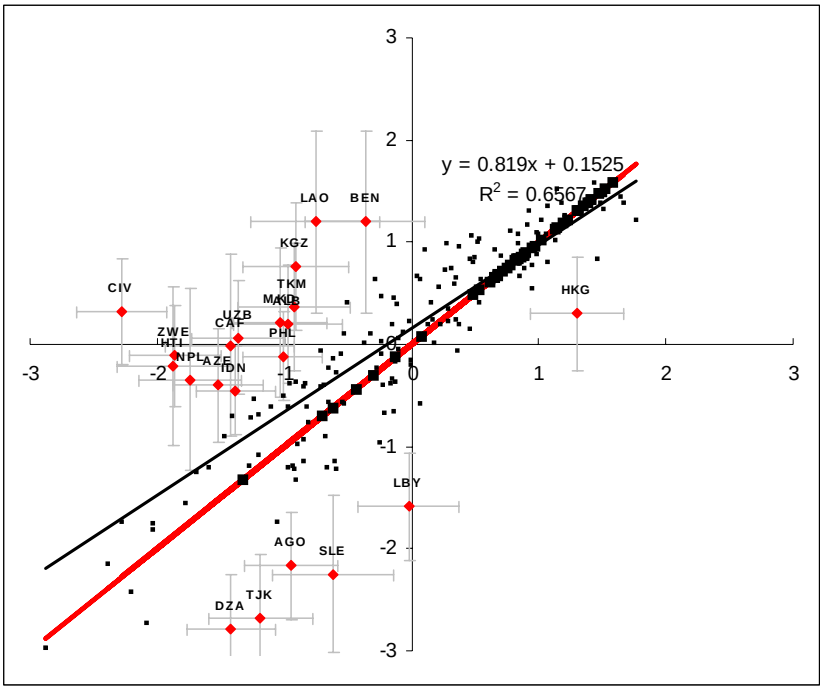


Figure 2: Changes Over Time in Governance Indicators 1996-2004

Political Stability and Absence of Violence



Control of Corruption

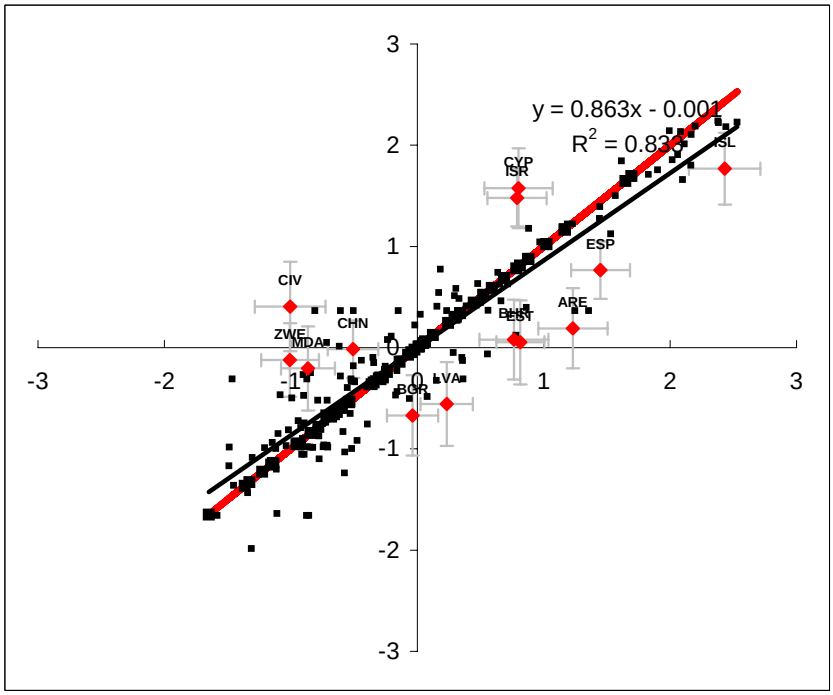


Figure 3: Significance of Changes in Individual Measures of Governance

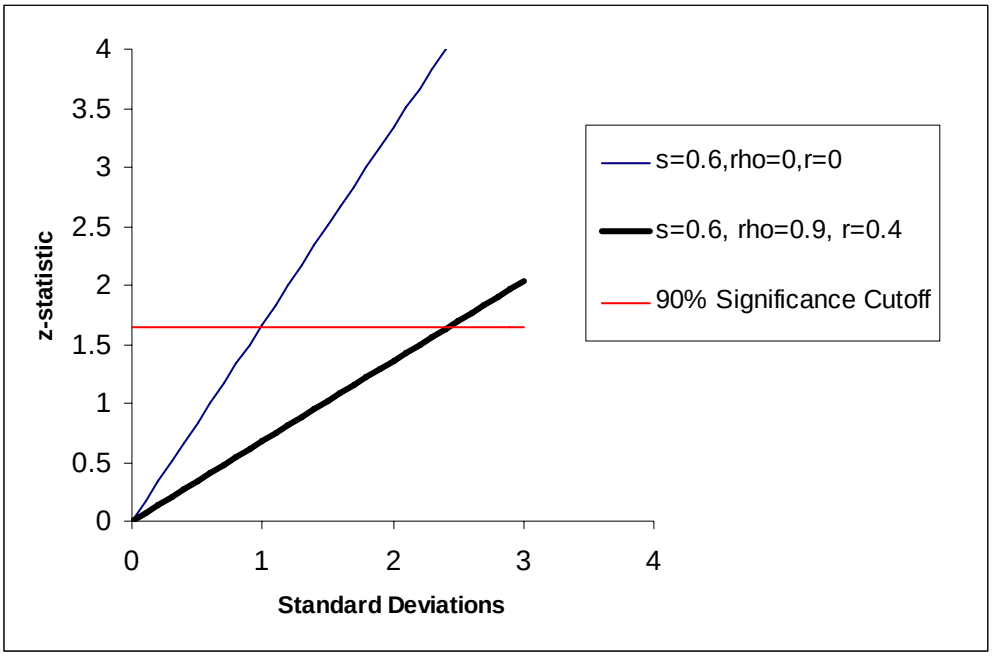


Figure 4: Changes in Measures of Ease of Business Entry, 2003-2004

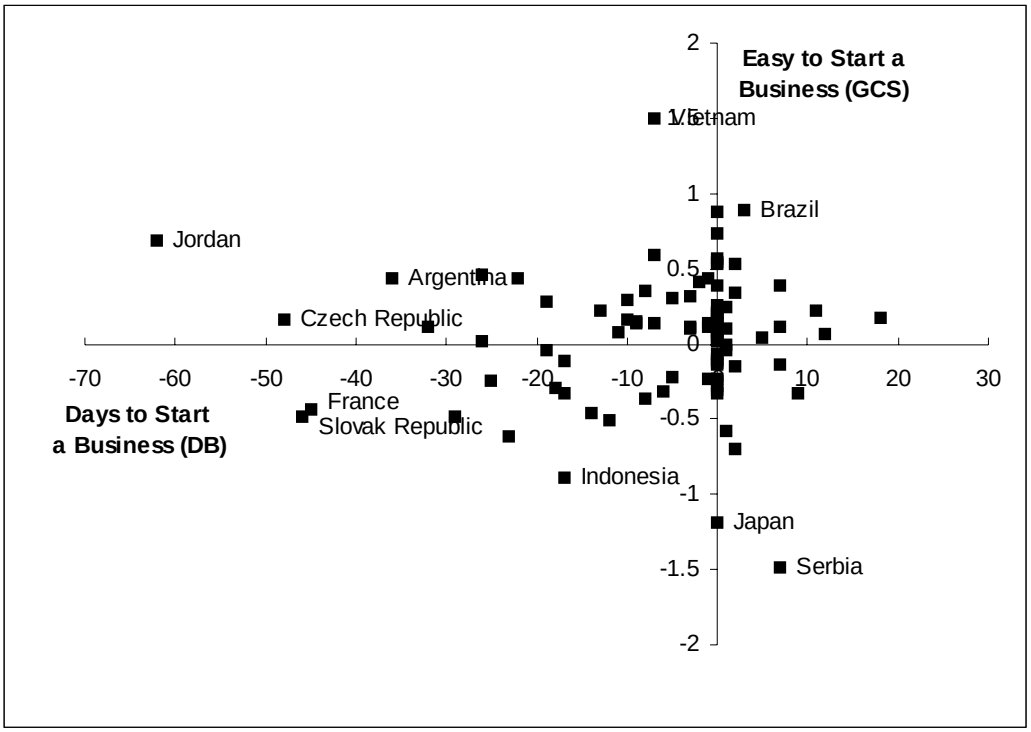


Figure 5: Halo Effects

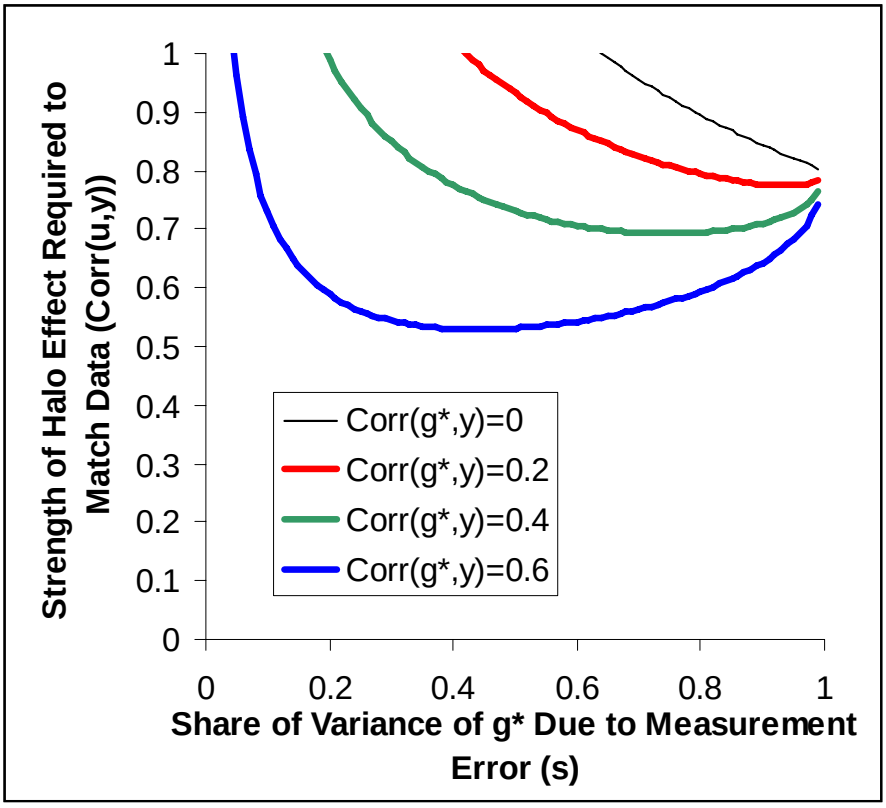


Figure 6: Governance and Per Capita Incomes in Africa

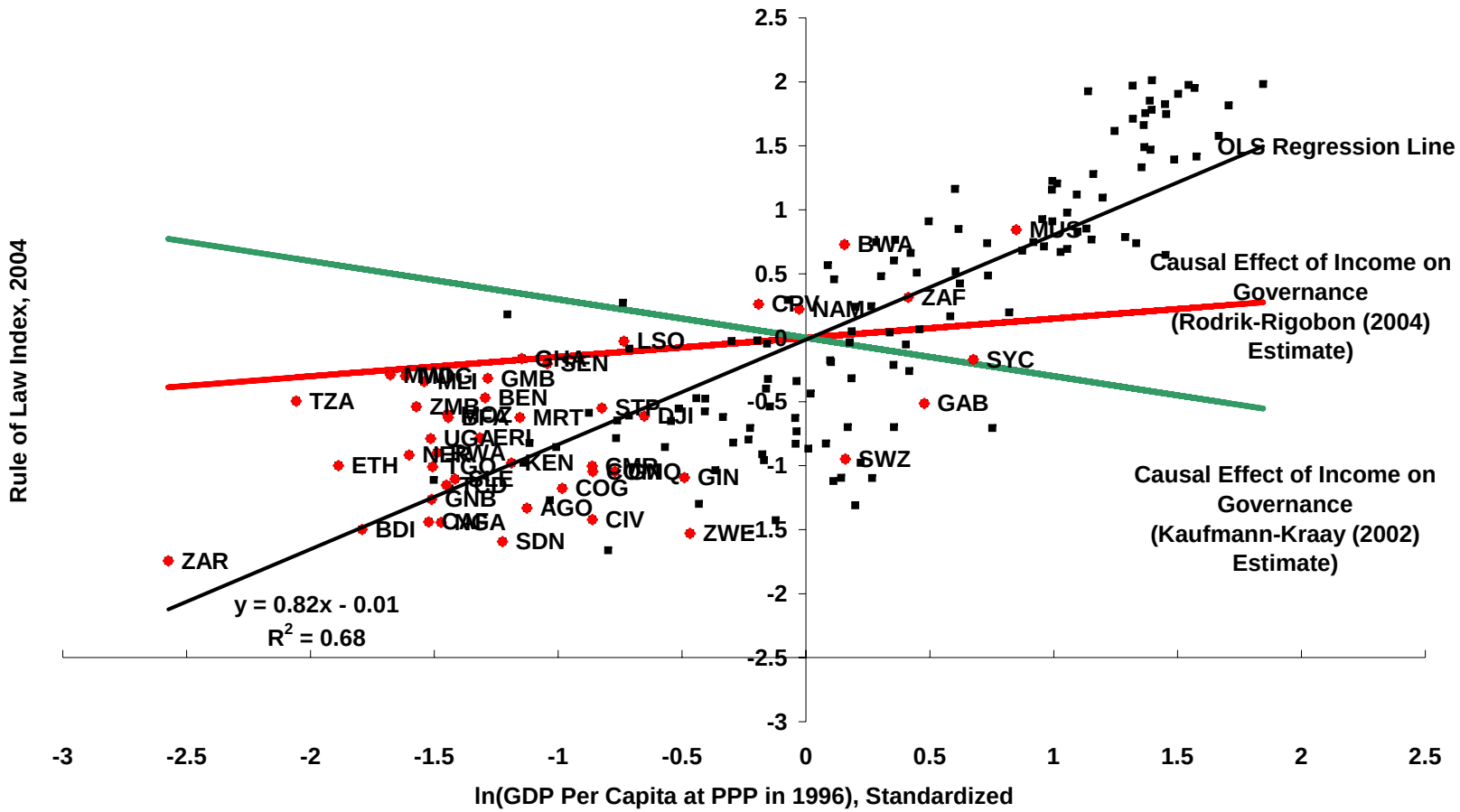


Table 1: Sources of Governance Data

Source	Publication	Code	Type 1/	Country Coverage 2/	Repre-sentative	1996	1998	2000	2002	2004
African Development Bank*	Country Policy & Institutional Assessments	ADB	P	50				x	x	x
Afrobarometer	Afrobarometer Survey	AFR	S	12					x	x
Asian Development Bank*	Country Policy & Institutional Assessments	ASD	P	26				x	x	x
Bertelsmann Foundation*	Bertelsmann Transformation Index	BTI	P	116						x
Brown University's Center for Public Policy*	Global E-Governance	EGV	P	192	x				x	x
Business Environment Risk Intelligence	Business Risk Service	BRI	P	50		x	x	x	x	x
Business Environment Risk Intelligence	Qualitative Risk Measure	QLM	P	115	x	x	x	x	x	x
Columbia University	State Capacity Project	CDU	P	98	x			x	x	x
Economist Intelligence Unit	Country Risk Service	EIU	P	115	x	x	x	x	x	x
European Bank for Reconstruction & Development	Transition Report	EBR	P	26		x	x	x	x	x
Freedom House*	Countries at the Crossroads	CCR	P	30						x
Freedom House	Nations in Transition	FHT	P	27		x	x	x	x	x
Freedom House	Freedom in the World	FRH	P	192	x	x	x	x	x	x
Furnar*	Index of Budget Transparency	LAI	S	10					x	x
Gallup International	Gallup Millennium Survey	GMS	S	60				x		
Gallup International	50th Anniversary Survey	GLP	S	44			x			
Gallup International	Voice of the People Survey	GAL	S	62					x	x
Global Insight's DRI McGraw-Hill	Country Risk Review	DRI	P	111	x	x	x	x	x	x
Heritage Foundation/Wallstreet Journal	Economic Freedom Index	HER	P	161	x	x	x	x	x	x
IJET Travel Intelligence*	Country Security Risk Assessment	IJT	P	167	x					x

1/ P=Poll, S=Survey
2/ Countries included in most recently available version of source
* indicates new source added in 2004

Table 1, Cont'd: Sources of Governance Data

Source	Publication	Code	Type 1/	Country Coverage 2/	Repre-sentative	1996	1998	2000	2002	2004
Institute for Management and Development	World Competitiveness Yearbook	WCY	S	49		x	x	x	x	x
International Research & Exchanges Board*	Media Sustainability Index	MSI	P	18					x	x
Latinobarometro	Latinobarometro Surveys	LBO	S	17		x	x	x	x	x
Merchant International Group*	Gray Area Dynamics	MIG	P	155	x				x	x
Political & Economic Risk Consultancy*	Corruption Survey	PRC	S	14			x	x	x	x
Political Risk Services	International Country Risk Guide	PRS	P	140	x	x	x	x	x	x
PriceWaterhouseCoopers	Opacity Index	PWC	S	35				x		
Reporters Without Borders	Reporters Without Borders	RSF	P	138	x				x	x
State Department / Amnesty International	Human Rights Dataset	HUM / PTS	P	192	x	x	x	x	x	x
United Nations Economic Commission for Africa*	Africa Governance Indicators	AGI	P	23				x	x	x
USAID / Vanderbilt University*	Democracy Surveys in Central America	USD	S	8						x
World Bank	Business Enterprise Environment Survey	BPS	S	27				x	x	x
World Bank	World Business Environment Survey	WBS	S	80	x		x	x		
World Bank	Country Policy & Institutional Assessments	PIA	P	136		x	x	x	x	x
World Economic Forum	Global Competitiveness Report	GCS	S	104	x	x	x	x	x	x
World Economic Forum	Africa Competitiveness Report	GCSA	S	23			x			
World Markets Research Center	World Markets Online	WMO	P	202	x				x	x

1/ P=Poll, S=Survey
2/ Countries included in most recently available version of source
* indicates new source added in 2004

Table 2: Summary Statistics on Governance Indicators

	Voice and Accountability	Political Stability	Government Effectiveness	Regulatory Quality	Rule of Law	Control of Corruption	Average
Number of Countries							
1996	192	165	180	182	167	151	173
1998	192	166	184	185	186	184	183
2000	192	166	187	188	188	187	185
2002	199	186	202	197	197	197	196
2004	207	207	209	204	208	204	207
Median Number of Sources Per Country							
1996	4	4	4	4	6	4	4
1998	4	4	4	4	7	5	5
2000	5	6	6	5	8	7	6
2002	7	7	8	7	10	8	8
2004	8	8	9	8	11	8	9
Proportion of Countries with Only One Data Source							
1996	0.15	0.13	0.21	0.15	0.07	0.18	0.15
1998	0.14	0.10	0.19	0.13	0.11	0.18	0.14
2000	0.14	0.06	0.07	0.07	0.05	0.07	0.08
2002	0.10	0.10	0.05	0.07	0.07	0.08	0.08
2004	0.06	0.06	0.08	0.08	0.08	0.08	0.07
Average Standard Error							
1996	0.26	0.36	0.29	0.34	0.26	0.29	0.30
1998	0.25	0.30	0.29	0.34	0.25	0.25	0.28
2000	0.25	0.32	0.25	0.32	0.22	0.24	0.27
2002	0.21	0.26	0.22	0.22	0.19	0.20	0.22

Table 3: Classifying Countries for the MCA

	Control of Corruption			*WMO*	*DRI*	*GCS*
	2004	2000	1996	2004	2004	2004
Probability of Being Above the Median Is:						
			Number of Countries			
Below 10%	17	15	16	10	5	3
Below 25%	24	24	19	17	11	6
Between 25% and 75%	20	20	18	38	11	12
Above 75%	26	25	15	15	12	12
Above 90%	23	22	11	6	7	8
Total Number of Countries	70	69	52	70	34	30
			Proportion of Countries			
Below 10%	0.24	0.22	0.31	0.14	0.15	0.10
Below 25%	0.34	0.35	0.37	0.24	0.32	0.20
Between 25% and 75%	0.29	0.29	0.35	0.54	0.32	0.40
Above 75%	0.37	0.36	0.29	0.21	0.35	0.40
Above 90%	0.33	0.32	0.21	0.09	0.21	0.27
Average Standard Error	0.18	0.25	0.35	0.41	0.42	0.44

Table 4: Large Changes in Governance, 1996-2004

		Governance Score			Sources available in both periods				Sources Added			Weights	
		2004	1996	Change	Agree	No change	Dis-agree	Agree/ (agree+ Disagree)		2004 (balanced sources)	2004 (sources added)	Balanced	Added
Voice & Accountability													
BIH	BOSNIA-HERZEGOVINA	-0.14	-1.20	1.07	2	0	0	1.00	8	-0.19	-0.11	0.27	0.74
CAF	CENTRAL AFRICAN REPUBLIC	-1.20	-0.17	-1.03	2	0	0	1.00	3	-1.08	-1.24	0.67	0.36
HRV	CROATIA	0.46	-0.50	0.96	4	0	0	1.00	7	0.34	0.72	0.74	0.27
ERI	ERITREA	-1.96	-1.10	-0.86	1	1	0	1.00	4	-1.75	-2.03	0.62	0.41
GMB	GAMBIA	-0.59	-1.34	0.75	3	0	0	1.00	5	-0.46	-0.74	0.65	0.38
GHA	GHANA	0.39	-0.35	0.74	4	0	0	1.00	8	0.38	0.36	0.61	0.41
HTI	HAITI	-1.50	-0.46	-1.03	3	0	0	1.00	5	-1.66	-1.18	0.53	0.49
IDN	INDONESIA	-0.44	-1.15	0.71	4	0	1	0.80	7	-0.33	-0.53	0.58	0.44
ISR	ISRAEL	0.46	1.07	-0.62	4	0	1	0.80	5	0.45	0.38	0.78	0.24
CIV	IVORY COAST	-1.46	-0.19	-1.27	4	0	0	1.00	4	-1.34	-1.54	0.71	0.32
KGZ	KYRGYZ REPUBLIC	-1.06	-0.48	-0.58	3	0	0	1.00	6	-1.03	-1.02	0.66	0.35
MEX	MEXICO	0.36	-0.23	0.59	5	0	1	0.83	8	0.42	0.21	0.64	0.38
NPL	NEPAL	-1.00	0.14	-1.13	2	0	0	1.00	6	-0.87	-1.02	0.44	0.58
NGA	NIGERIA	-0.65	-1.49	0.84	4	0	0	1.00	9	-0.71	-0.54	0.52	0.50
PER	PERU	-0.04	-0.73	0.69	3	0	2	0.60	7	0.07	-0.23	0.65	0.37
SLE	SIERRA LEONE	-0.49	-1.37	0.88	2	0	1	0.67	5	-0.29	-0.67	0.53	0.49
SVK	SLOVAK REPUBLIC	1.10	0.37	0.72	5	0	0	1.00	5	1.06	1.04	0.86	0.16
YUG	SERBIA AND MONTENEGRO	0.12	-1.38	1.50	4	0	0	1.00	7	-0.06	0.22	0.36	0.65
ZWE	ZIMBABWE	-1.48	-0.30	-1.18	4	0	0	1.00	6	-1.60	-1.21	0.55	0.47
Political Stability													
ALB	ALBANIA	-0.97	0.20	-1.17	2	0	1	0.67	3	-0.60	-1.19	0.59	0.48
DZA	ALGERIA	-1.42	-2.78	1.36	3	0	1	0.75	5	-1.18	-1.46	0.58	0.46
AGO	ANGOLA	-0.95	-2.17	1.22	4	0	0	1.00	4	-0.73	-1.07	0.63	0.42
AZE	AZERBAIJAN	-1.52	-0.40	-1.12	2	0	1	0.67	5	-1.46	-1.32	0.53	0.52
BEN	BENIN	-0.37	1.20	-1.56	1	0	0	1.00	4	0.23	-0.52	0.28	0.80
CAF	CENTRAL AFRICAN REPUBLIC	-1.43	-0.01	-1.42	1	0	0	1.00	3	-1.13	-1.30	0.33	0.76
HTI	HAITI	-1.87	-0.21	-1.66	2	0	0	1.00	4	-1.06	-1.98	0.36	0.71
HKG	HONG KONG	1.30	0.30	1.00	3	0	1	0.75	5	0.80	1.60	0.54	0.51
IDN	INDONESIA	-1.38	-0.45	-0.93	6	0	0	1.00	6	-1.33	-1.20	0.60	0.44
CIV	IVORY COAST	-2.28	0.32	-2.60	4	0	0	1.00	4	-2.11	-2.04	0.59	0.45
KGZ	KYRGYZ REPUBLIC	-0.91	0.76	-1.68	2	0	0	1.00	4	-0.77	-0.86	0.45	0.61
LAO	LAOS	-0.76	1.20	-1.95	1	0	0	1.00	3	-0.99	-0.51	0.33	0.76
LBY	LIBYA	-0.02	-1.59	1.57	3	0	1	0.75	2	-0.13	0.26	0.75	0.30
MKD	MACEDONIA	-1.04	0.21	-1.25	1	1	0	1.00	6	-0.75	-1.06	0.39	0.66
NPL	NEPAL	-1.74	-0.35	-1.39	1	0	0	1.00	5	-1.40	-1.61	0.21	0.85
PHL	PHILIPPINES	-1.01	-0.12	-0.90	4	0	2	0.67	6	-0.80	-1.13	0.60	0.44
SLE	SIERRA LEONE	-0.61	-2.25	1.64	2	0	0	1.00	3	-0.10	-0.81	0.41	0.67
TJK	TAJIKISTAN	-1.19	-2.67	1.48	2	0	0	1.00	4	-0.91	-1.19	0.45	0.61
TKM	TURKMENISTAN	-0.92	0.36	-1.29	2	0	0	1.00	3	-1.22	-0.47	0.51	0.56
UZB	UZBEKISTAN	-1.37	0.07	-1.43	3	0	0	1.00	4	-1.74	-0.67	0.56	0.48
ZWE	ZIMBABWE	-1.86	-0.11	-1.74	4	0	1	0.80	3	-1.51	-2.00	0.66	0.39
Government Effectiveness													
ARG	ARGENTINA	-0.33	0.45	-0.78	5	1	1	0.83	6	-0.30	-0.36	0.66	0.37
CIV	IVORY COAST	-1.30	-0.11	-1.19	4	0	0	1.00	6	-1.21	-1.24	0.51	0.52
LVA	LATVIA	0.60	0.04	0.56	3	0	1	0.75	7	0.63	0.40	0.63	0.39
LTU	LITHUANIA	0.70	0.06	0.64	3	1	0	1.00	8	0.65	0.63	0.59	0.43
SLE	SIERRA LEONE	-1.32	-0.24	-1.07	2	0	0	1.00	5	-0.83	-1.38	0.30	0.74
TZA	TANZANIA	-0.37	-1.18	0.81	4	0	0	1.00	9	-0.36	-0.37	0.35	0.67
ZWE	ZIMBABWE	-1.20	-0.26	-0.94	4	0	1	0.80	6	-0.95	-1.35	0.54	0.49

Table 4, Cont'd: Large Changes in Governance, 1996-2004

		Governance Score			Sources available in both periods				Sources Added			Weights	
		2004	1996	Change	Agree	No change	Dis-agree	Agree/ (agree+ Disagree)		2004 (balanced sources)	2004 (sources added)	Balanced	Added
Regulatory Quality													
ARG	ARGENTINA	-0.81	0.82	-1.63	7	0	0	1.00	3	-0.87	-0.44	0.79	0.25
BOL	BOLIVIA	0.05	0.82	-0.77	3	0	2	0.60	4	0.25	-0.53	0.72	0.32
BIH	BOSNIA-HERZEGOVINA	-0.66	-2.09	1.43	1	0	0	1.00	8	-0.72	-0.60	0.24	0.80
CUB	CUBA	-1.81	-0.77	-1.04	3	0	1	0.75	3	-1.89	-1.11	0.73	0.31
ISL	ICELAND	1.82	0.53	1.29	3	0	0	1.00	3	1.39	1.70	0.50	0.57
IDN	INDONESIA	-0.42	0.27	-0.69	3	0	4	0.43	4	-0.38	-0.49	0.77	0.26
LTU	LITHUANIA	1.16	0.38	0.79	4	0	1	0.80	6	0.99	1.07	0.67	0.37
MMR	MYANMAR	-2.34	-1.12	-1.23	3	0	1	0.75	3	-2.33	-1.68	0.73	0.31
PRY	PARAGUAY	-0.60	0.58	-1.18	3	0	1	0.75	4	-0.59	-0.56	0.70	0.34
SVK	SLOVAK REPUBLIC	1.15	0.27	0.88	6	0	1	0.86	5	1.03	0.97	0.75	0.28
VEN	VENEZUELA	-1.24	-0.08	-1.16	6	0	1	0.86	3	-1.29	-0.73	0.79	0.25
ZMB	ZAMBIA	-0.49	0.27	-0.76	2	0	3	0.40	6	-0.58	-0.33	0.62	0.41
ZWE	ZIMBABWE	-2.15	-0.87	-1.28	4	0	2	0.67	5	-2.21	-1.57	0.67	0.36
Rule of Law													
ARG	ARGENTINA	-0.71	0.28	-0.99	10	0	0	1.00	5	-0.71	-0.63	0.71	0.30
BRB	BARBADOS	1.21	-0.28	1.49	1	0	0	1.00	3	1.39	0.93	0.40	0.65
HRV	CROATIA	0.07	-0.53	0.60	4	1	1	0.80	7	-0.04	0.21	0.65	0.37
EST	ESTONIA	0.91	0.35	0.56	4	1	1	0.80	9	0.80	0.94	0.59	0.43
IDN	INDONESIA	-0.91	-0.36	-0.55	6	2	1	0.86	7	-0.79	-1.08	0.68	0.33
CIV	IVORY COAST	-1.42	-0.69	-0.74	5	0	0	1.00	6	-1.29	-1.45	0.57	0.45
LTU	LITHUANIA	0.60	-0.15	0.75	5	1	0	1.00	8	0.50	0.68	0.63	0.39
MLT	MALTA	1.23	0.04	1.18	1	0	1	0.50	4	1.01	1.20	0.37	0.67
PRY	PARAGUAY	-1.09	-0.50	-0.59	4	2	0	1.00	6	-0.96	-1.16	0.54	0.47
PHL	PHILIPPINES	-0.62	-0.11	-0.50	6	2	1	0.86	5	-0.56	-0.70	0.71	0.31
SAU	SAUDI ARABIA	0.20	0.75	-0.56	3	2	2	0.60	4	0.36	-0.22	0.70	0.31
SWZ	SWAZILAND	-0.95	0.40	-1.34	2	0	0	1.00	6	-0.42	-1.06	0.25	0.77
THA	THAILAND	-0.05	0.49	-0.54	6	1	2	0.75	5	-0.10	0.02	0.71	0.31
ZWE	ZIMBABWE	-1.53	-0.24	-1.29	7	0	0	1.00	7	-1.40	-1.58	0.62	0.40
Control of Corruption													
BHR	BAHRAIN	0.76	0.08	0.68	2	0	2	0.50	4	0.57	0.93	0.67	0.36
BGR	BULGARIA	-0.04	-0.67	0.63	3	0	1	0.75	8	-0.11	-0.02	0.37	0.65
CHN	CHINA	-0.51	-0.01	-0.49	4	1	2	0.67	5	-0.66	-0.31	0.56	0.46
CYP	CYPRUS	0.80	1.58	-0.77	3	0	1	0.75	3	0.61	0.96	0.66	0.36
EST	ESTONIA	0.82	0.05	0.76	3	0	0	1.00	10	0.58	0.83	0.30	0.72
ISL	ICELAND	2.43	1.77	0.66	1	1	2	0.33	3	2.22	2.33	0.72	0.30
ISR	ISRAEL	0.79	1.48	-0.69	4	0	2	0.67	4	0.81	0.48	0.75	0.27
CIV	IVORY COAST	-1.01	0.41	-1.41	3	0	0	1.00	5	-0.88	-1.02	0.44	0.58
LVA	LATVIA	0.23	-0.56	0.79	3	0	0	1.00	8	0.24	0.18	0.35	0.67
MDA	MOLDOVA	-0.86	-0.21	-0.66	2	0	1	0.67	7	-0.90	-0.80	0.38	0.64
ESP	SPAIN	1.45	0.77	0.68	5	2	0	1.00	4	1.42	1.16	0.75	0.27
ARE	UNITED ARAB EMIRATES	1.23	0.19	1.04	4	0	0	1.00	3	1.15	1.10	0.67	0.36
ZWE	ZIMBABWE	-1.01	-0.12	-0.89	4	0	1	0.80	6	-0.92	-1.02	0.55	0.47

Table 5: Agreement Ratio for All Changes in Governance, 1996-2004

ALL CHANGES

	Sample	Agree	No Change	Disagree	Agree / (Agree + Disagree)
Voice and Accountability	192	1.50	0.52	0.80	0.65
Political Stability	165	1.58	0.22	0.69	0.70
Government Effectiveness	180	1.51	0.51	0.70	0.68
Regulatory Quality	182	1.74	0.13	1.11	0.61
Rule of Law	167	1.62	1.34	1.13	0.59
Control of Corruption	151	1.12	0.86	0.67	0.63
Average	**173**	**1.51**	**0.60**	**0.85**	**0.64**

SIGNIFICANT CHANGES (90%)

	Sample	Agree	No Change	Disagree	Agree / (Agree + Disagree)
Voice and Accountability	19	3.32	0.05	0.32	0.93
Political Stability	21	2.52	0.05	0.38	0.91
Government Effectiveness	17	3.57	0.29	0.43	0.91
Regulatory Quality	13	3.69	0.00	1.31	0.76
Rule of Law	14	4.57	0.86	0.64	0.87
Control of Corruption	13	3.15	0.31	0.92	0.78
Average	**16**	**3.47**	**0.26**	**0.67**	**0.86**

Table 6: Global Trends in Governance 1996-2004 for Selected Sources

[Quasi-Balanced Sample]* ** * †**

	# of Countries	World Average					Std. Dev. Across Countries					t-statistic for mean difference 1996-2004
		1996	1998	2000	2002	2004	1996	1998	2000	2002	2004	
Voice and Accountability												
EIU	115	0.41	0.42	0.42	0.46	0.46	*0.30*	*0.32*	*0.31*	*0.28*	*0.28*	1.5
PRS *	140	0.63	0.63	0.63	0.63	0.65	*0.25*	*0.27*	*0.27*	*0.26*	*0.26*	0.7
GCS **	88	..	..	..	0.49	0.51	..	..	..	*0.14*	*0.14*	….
FRH (PR+CL)	190	0.56	0.58	0.58	0.60	0.62	*0.34*	*0.33*	*0.33*	*0.33*	*0.33*	1.7
FRH (Press Freedom)	188	0.54	0.53	0.54	0.55	0.53	*0.24*	*0.26*	*0.25*	*0.25*	*0.25*	-0.2
WMO	186	..	..	..	0.55	0.53	..	..	..	*0.26*	*0.22*	…
Political Stability												
EIU	115	0.54	0.51	0.56	0.54	0.56	*0.29*	*0.30*	*0.30*	*0.28*	*0.26*	0.7
PRS *	140	0.78	0.71	0.73	0.75	0.75	*0.15*	*0.20*	*0.17*	*0.14*	*0.13*	-1.5
GCS **	88	0.73	0.74	0.63	0.66	0.67	*0.14*	*0.16*	*0.18*	*0.17*	*0.13*	-2.5 [-2.4]†
WMO	186	..	..	..	0.67	0.56	..	..	..	*0.24*	*0.20*	…
Government Effectiveness												
EIU	115	0.39	0.45	0.44	0.38	0.38	*0.30*	*0.24*	*0.24*	*0.29*	*0.30*	-0.2
PRS *	140	0.63	0.66	0.57	0.64	0.62	*0.19*	*0.12*	*0.11*	*0.17*	*0.15*	-0.4
GCS **	82	0.48	0.52	0.53	0.54	0.55	*0.22*	*0.28*	*0.27*	*0.24*	*0.23*	1.9 [2.8]†
WMO	186	..	..	..	0.56	0.55	..	..	..	*0.23*	*0.22*	…
Regulatory Quality												
EIU	115	0.42	..	..	0.51	0.55	*0.25*	..	..	*0.25*	*0.23*	4.3
GCS **	82	0.43	0.44	0.38	0.31	0.35	*0.15*	*0.15*	*0.16*	*0.14*	*0.13*	-3.4 [-3.0]†
WMO	186	..	..	..	0.58	0.61	..	..	..	*0.25*	*0.17*	…
HERITAGE ***	155	0.50	0.48	0.49	0.50	0.50	*0.30*	*0.31*	*0.31*	*0.29*	*0.28*	0.0
Rule of Law												
EIU	115	0.47	0.50	0.48	0.52	0.52	*0.27*	*0.30*	*0.30*	*0.26*	*0.26*	1.4
PRS *	140	0.72	0.62	0.65	0.62	0.63	*0.23*	*0.26*	*0.23*	*0.24*	*0.22*	-3.4
GCS **	82	0.67	0.63	0.57	0.51	0.51	*0.18*	*0.24*	*0.25*	*0.22*	*0.22*	-4.6 [-2.9]†
WMO	186	..	..	..	0.58	0.57	..	..	..	*0.23*	*0.20*	…
HERITAGE ***	155	0.50	0.48	0.46	0.44	0.44	*0.30*	*0.31*	*0.31*	*0.29*	*0.28*	-1.8
QLM	115	0.45	0.45	0.46	0.46	0.45	*0.29*	*0.29*	*0.30*	*0.30*	*0.30*	0.1
Control of Corruption												
EIU	115	0.35	0.34	0.33	0.35	0.35	*0.31*	*0.32*	*0.31*	*0.32*	*0.33*	0.2
PRS *	140	0.59	0.51	0.47	0.41	0.42	*0.21*	*0.21*	*0.21*	*0.19*	*0.19*	-7.2
GCS **	82	..	0.66	0.69	0.64	0.66	..	*0.29*	*0.25*	*0.22*	*0.21*	0.0 [-0.1]†
WMO	186	..	..	..	0.52	0.54	..	..	..	*0.27*	*0.20*	…
QLM	115	0.39	0.40	0.40	0.40	0.38	*0.29*	*0.29*	*0.29*	*0.29*	*0.29*	-0.2

Note that all variables are scaled to run from 0 to 1

* PRS Country coverage in 1996: 129 countries, all other periods 140.

** GCS Country coverage in 1996: 58; in 1998: 59; in 2000: 75; and in 2002 and in 2004: 82.

*** Heritage Country coverage in 1996: 137; all other periods 155.

† Values in square brackets for GCS report t-stats for fully balanced sample from 1996 (same 52 countries)

Table 7: Persistence and Inference About Changes in Governance Over Time

Summary Statistics

	Correlations			Mean Absolute Changes		Persistence	
	Levels, 2004	Levels, 1996	Changes, 1996-2004	Static	Dynamic	Governance	Average for Source Errors
VA	1.00	0.99	0.96	0.27	0.14	0.93	0.39
PV	0.99	0.99	0.98	0.44	0.30	0.78	0.39
GE	0.99	0.99	0.93	0.27	0.11	0.92	0.35
RQ	0.99	0.99	0.93	0.36	0.21	0.86	0.36
RL	0.99	0.99	0.88	0.23	0.12	0.94	0.53
CC	0.99	0.99	0.92	0.33	0.16	0.89	0.50
Average	0.99	0.99	0.93	0.32	0.17	0.89	0.42

Consequences of Persistence for Inference

	Mean t-Statistics		Number Significant at 90%		Rule of Thumb	
	Static	Dynamic	Static	Dynamic	Number Significant	Also Significant in Dynamic Model
VA	0.85	0.57	26	13	12	12
PV	0.91	0.78	21	18	14	14
GE	0.69	0.41	12	1	1	1
RQ	0.86	0.63	25	14	11	9
RL	0.73	0.55	16	7	7	5
CC	0.90	0.58	26	7	10	7
Average	0.82	0.59	21	10	9	8

Table 8: De Jure and De Facto Measures

Dependent Variable is GCS '04: "Easy to Start a Business?"

	1	2	3	4	5
All Countries					
# of Days to start business (DB '04)	-1.18	-0.43	-0.47	-0.60	-0.59
	5.46***	1.87*	1.96*	4.33***	4.19***
Corporate Tax Rate			-0.01		0.01
			1.06		0.69
Control of Corruption (2002)		0.47	0.45	0.18	0.18
		6.14***	5.84***	2.80***	2.81***
Administrative Regulations (GCS '04)				0.75	0.77
				9.86***	9.05***
Observations (# of countries)	**81**	**81**	**81**	**81**	**81**
Adjusted R-squared	**0.23**	**0.44**	**0.44**	**0.71**	**0.71**
Developing Countries					
# of Days to start business (DB '04)	-0.49	-0.32	-0.29	-0.49	-0.47
	1.44	0.95	0.86	2.42**	2.25**
Corporate Tax Rate			0.01		0.01
			0.66		0.73
Control of Corruption (2002)		0.50	0.53	0.19	0.22
		3.30***	3.08***	1.48	1.67
Administrative Regulations (GCS '04)				0.83	0.82
				8.76***	8.73***
Observations (# of countries)	**56**	**56**	**56**	**56**	**56**
Adjusted R-squared	**0.01**	**0.19**	**0.18**	**0.57**	**0.57**
OECD + Newly-Industrialized Countries					
# of Days to start business (DB '04)	-0.97	-0.53	-0.57	-0.73	-0.74
	3.29***	1.65	1.88*	3.41***	3.33***
Corporate Tax Rate			-0.04		0.00
			1.92*		0.09
Control of Corruption (2002)		0.75	0.62	0.29	0.29
		2.85***	2.38**	1.28	1.25
Administrative Regulations (GCS '04)				0.64	0.65
				4.44***	3.51***
Observations (# of countries)	**25**	**25**	**25**	**25**	**25**
Adjusted R-squared	**0.18**	**0.36**	**0.46**	**0.69**	**0.67**

Note: DB refers to "Doing Business" study, GCS refers to Global Competitiveness Survey

Table 8, Cont'd: De Jure and De Facto Measures

Dependent Variable is GCS '04: "How Heavy Is Overall Tax Burden?"

	1	2	3	4	5
All Countries					
# of Days to start business (DB '04)			-0.96		-0.27
			0.46		0.15
Corporate Tax Rate	0.29	0.28	0.27	0.18	0.18
	2.37**	2.29**	2.22**	1.58	1.55
Control of Corruption (2002)		-0.77	-0.96	0.58	0.52
		1.27	1.19	0.91	0.62
Administrative Regulations (GCS '04)				-4.29	-4.28
				3.91***	3.91***
Observations (# of countries)	**81**	**81**	**81**	**81**	**81**
Adjusted R-squared	**0.09**	**0.09**	**0.08**	**0.24**	**0.23**
Developing Countries					
# of Days to start business (DB '04)			-2.06		-1.46
			0.68		0.54
Corporate Tax Rate	0.11	0.02	0.01	0.03	0.02
	0.71	0.15	0.09	0.16	0.11
Control of Corruption (2002)		-2.66	-2.80	-1.59	-1.71
		1.78*	1.88*	1.07	1.16
Administrative Regulations (GCS '04)				-2.93	-2.87
				1.62	1.60
Observations (# of countries)	**56**	**56**	**56**	**56**	**56**
Adjusted R-squared	**0.00**	**0.04**	**0.03**	**0.09**	**0.08**
OECD + Newly-Industrialized Countries					
# of Days to start business (DB '04)			0.96		2.37
			0.35		0.93
Corporate Tax Rate	0.63	0.64	0.64	0.33	0.32
	4.59***	4.43***	4.52***	3.56***	3.90***
Control of Corruption (2002)		0.47	0.78	2.63	3.49
		0.23	0.32	1.70	1.94*
Administrative Regulations (GCS '04)				-5.15	-5.38
				4.54***	4.89***
Observations (# of countries)	**25**	**25**	**25**	**25**	**25**
Adjusted R-squared	**0.47**	**0.45**	**0.42**	**0.65**	**0.65**

Note: DB refers to "Doing Business" study, GCS refers to Global Competitiveness Survey