[file: g242004UNCTAD]
03 March 2004

DRAFT
XVIII Technical Group
Meeting of the Group of 24
Geneva, March 8-9, 2004

# HOW WELL DO MEASUREMENTS OF AN ENABLING DOMESTIC ENVIRONMENT FOR DEVELOPMENT STAND UP?

**Barry Herman**[1]

***Abstract***

Official donors and private investors, focused increasingly on the role of institutions and policy quality in development, seek to monitor them more closely. This paper examines the World Bank's Country Policy and Institutional Assessments, the World Economic Forum's Global Competitiveness Indices, and a set of governance indicators also developed at the World Bank. The paper calls for appreciating the weaknesses of such indicators, especially their low ability to discriminate among countries or over time. More robust elements in such indicators, however, may usefully complement structured narrative analyses of countries and stimulate public discourse on institutional and policy development.

*Keywords*: World Bank, International Development Association, Country Policy and Institutional Assessment, Global Competitiveness Index, governance, enabling environment, selectivity, official development assistance.

[1] Formerly, Chief, Policy Analysis and Development Branch, Financing for Development Office, Department of Economic and Social Affairs, United Nations, New York (contact: herman@un.org). Views expressed are those of the author and not necessarily the views of the United Nations. Research assistance on the CPIA by Daniel Cohen is much appreciated. We both express our appreciation to Rui Coutinho of the World Bank for patiently explaining how the CPIA process works. Appreciation is also due to Pingfan Hong for econometric consultations and to Nancy Alexander, Ariel Buira, Paul Collier, Tim Kessler, Andrew Kuper, Agosto Lopez-Claros, Cristián Ossa, Ejeviome Otobo, and Shang-Jin Wei for reading earlier drafts. If the paper is not much better as a result, it is my fault.

*"It's the institutions, stupid."*
Guillermo Calvo and Frederic Mishkin[2]

There is a great sense among economists working on developing countries today that the quality and robustness of domestic political and economic institutions matter greatly, both for the effectiveness of all types of policies (including exchange rate management, the focus of Calvo and Mishkin in the quote above) and for the prospects for development itself. In this view, if societies get their institutions "right" and also adopt the "right" policies (which is supposed to be more likely when a country has the "right" institutions), they will create an "enabling environment" for development that will transform positive economic stimuli into long-lived, virtuous circles of development.

Official donor and creditor agencies thus proclaim to their developing and transition economy clients the absolute necessity of building an "enabling" domestic environment for development. The phrase is a social science term of art in that it defines itself by its results rather than its characteristics. However, a number of specific domestic policies and institutions have been especially important candidates for the list of required qualities of an enabling environment because they are the ones advocated by the Bretton Woods institutions.[3]

Progress in implementing these policies is assessed today by the World Bank in its Country Policy and Institutional Assessment (CPIA) indicators. Different institutions attempt to capture partly overlapping sets of "essential" features in other indicators, such as those by the World Economic Forum (WEF). In addition, some authors have combined the indicators developed by various others into statistically derived synthetic measures, notably a team at the World Bank led by Daniel Kaufmann.

The remainder of this paper critically examines the World Bank's CPIA methodology, focusing on the CPIA exercise for 2003. It is contrasted with the two main annual indicator exercises of the WEF. Finally, the approach undertaken by the Kaufmann team is discussed. The aim is to gain insight into how such indicators are constructed and some notion of their reliability. The conclusion is that all such indicators should claim no more than to be windows into a partial and clouded picture of development. One should be wary of asking the indicators to do policy jobs for which they are weakly suited (admittedly, this begs the question of how such policy jobs should be done, but that is beyond the scope of this paper). Indeed, to motivate the more technical discussion, the following section tries to set the context for why these indicators matter.

---

[2] Calvo and Mishkin, 2003, p. 106.

[3] A somewhat more eclectic set of principles than were usually advocated was adopted in the Monterrey Consensus of the International Conference on Financing for Development in 2002 by a more inclusive process and grouping of decision makers than drive the decision making in Washington (see United Nations, 2002, pp. 4-5).

## Development cooperation and development advocacy: why the indicators matter

For over 25 years, the World Bank has sought to measure progress in implementation of the Bretton Woods set of necessary policies and institutions for an enabling domestic environment for development, albeit under an evolving perception of what to include within its scope. The chief motivation for the measurement has from the beginning been the directive of the Bank's Governors that the Bank should allocate the concessional resources of its International Development Association (IDA) according to the "policy performance" of borrowing countries, as well as their need. As the requirement to measure progress was interpreted quantitatively, the Bank developed a set of numerical indicators that were scored in personal judgements by the staff that dealt with the relevant policies in the countries concerned. For most of the last quarter century, the quantitative assessment exercise was fully confidential (although World Bank researchers were allowed to use the indicators on condition that confidentiality was maintained, making replication by other scholars impossible). However, beginning in 2000 the Bank revealed its methodology and summary statistics of its measurements for groups of countries and for groups of components of the overall CPIA indicator. It also began to show individual countries their CPIA scores.

A major controversy in the Bank at the current moment is whether the CPIA scores for specific policy and institutional areas and for individual countries should also be brought into the public domain. This paper argues they should not because the methodology is too weak and unreliable for the scores to merit the attention they might receive if published. Indeed, greater confidence in the objectivity of CPIA-type assessments might be warranted were they undertaken by independent, country-based scholars and not Bank staff.

The proper place for Bank staff assessments — as equally for those of account executives in private financial institutions — is in confidential reports to help management make the decisions for which it is responsible. The CPIA was originally developed to serve such a management function, but it is increasingly asked to serve a broader social and political analysis function, assessing whether countries are actually building an (not "the") enabling environment. Moreover, as the aid community increasingly adopts "selectivity" criteria in deciding which countries to support and which reforms to promote through policy conditionality (see Koeberle, 2003, pp. 257-260), the weight that might be placed on a set of publicly available CPIA scores could exceed what they should bear.[4] Indeed, if allocations to countries from the new Millennium Challenge Account of the United States were to be based on CPIA ratings, they would apparently differ significantly from what is likely based on the indicators the United States is expected to use (see Alexander 2004, p. 8).

In fact, numerous independent organizations and institutions have developed indicators of institutional and policy "quality", although most of them seek to measure a

---

[4] See also Hout (2003) on implications of such indicators for "selectivity" in assistance, in particular regarding IDA, Dutch assistance and the Millennium Challenge Account of the United States.

rather narrow range of aspects of what their authors take to be the “enabling” environment for development (e.g., Transparency International on perceptions of corruption, Freedom House on political rights and civil liberties, PricewaterhouseCoopers on opacity in private and government policies and reporting). These indicators are transparent in their methodology and the author institutions make their results fully available (although some business-oriented indicators are sold to paying clients instead of being freely provided).

One prominent independent effort to produce a comprehensive indicator like the CPIA is the World Economic Forum’s (WEF) two Global Competitiveness indices. Like most of the partial indicators, as noted above, WEF aims to focus public attention on individual country performance in order to raise public debate about shortcomings, as it measures them. For sure, the developers of the WEF indicators have a particular perspective and the reader needs to be aware of what it is. However, that is not hidden from view, as the whole exercise hinges on transparency. Indeed, while this paper takes issue with some of the methodological decisions in the index, it can be said that the WEF does what it sets out to do, namely reflect the views on the “enabling” environment of the internationally oriented business community in poor and rich countries.

In addition to the need for transparency, one needs to appreciate that the indicators are quite weak in the sense that one should not put much faith in the precise numbers assigned by them. The best demonstration of this point seems to be in the conclusions coming from a research project developed by Daniel Kaufmann and colleagues at the World Bank that will be discussed below. They measure different parts of the domestic economic and political environment by drawing comprehensively on information from a large number of different exercises that attempt to measure various institutional and policy areas in individual countries. Some of the information is based on survey data, other parts are expert assessments like the CPIA and there are also “hard data” indicators. The authors statistically combine all the information into a series of synthetic indicators in a way that is rich, in being based on the most information feasible, and also yields measures of how much confidence is warranted in the results derived. Kaufmann’s results are sobering and should serve as a warning that individual measures — whether of the CPIA, the WEF indicators, or indicators of particular aspects of institutions or policy quality — are not able to discriminate reliably among countries except when their scores are quite far from each other (more precisely, the authors find 90% confidence intervals are wide). The same warning would apply to drawing conclusions from scores for the same country at different points in time.

Despite their weaknesses, there continues to be significant interest in such measures. They seem to speak to a need. They were, for example, considered potential material from the start for a new annual series of joint reports from the World Bank and International Monetary Fund on the implementation of policies and actions needed to reach international development goals such as are contained in the Millennium Declaration. The first report is expected to be available before the 2004 Spring Meetings of the Bretton Woods institutions and is expected to be a focus of discussion in the Development Committee. In discussions in 2003 of proposed approaches to the report, the Bank and Fund indicated that they would seek to capture the degree of relevant policy implementation by developing countries, donor countries and international institutions along a number of dimensions. For domestic policy in developing and transition economies, the first indicators are summary CPIA statistics. However, this paper argues that there are a number of rather curious features in the current CPIA methodology that should cast doubt on the meaning of the results. Happily, other indicators may also be utilized in the report.

Presumably, the object of the exercise being prepared for the Development Committee is to bring political pressure on the laggards and international support for the more advanced performers. As actual decision-making is at the level of individual countries and international institutions, the monitoring should also be at individual country or institution level. It should highlight the achievements and shortcomings of each developed as well as each developing and transition economy country, and the same applies to the reporting on multilateral institutions. In fact, this is not planned. Such reporting is not even possible for those elements of the "enabling" environment for which the CPIA components might serve as indicators, owing to the confidential nature of CPIA scores. The implication of this paper is that in light of the weakness of the CPIA itself, this should not be bemoaned. Many other quantitative indicators do not have this restriction and are already published, country-by-country and item-by-item.

In other words, the independent institutions that have been developing quantitative indicators of policy quality and institutional development of the countries of the world by and large seem to serve a useful public function in focusing attention on one country's performance compared to another (their methodologies, being open to scrutiny, can also be critiqued publicly). Yet, this type of comparison, even when done objectively and well, must be understood to be a gross indicator, whose importance lies in the public discussion of the numbers rather than in the numbers themselves. Indeed, the best methodology would seem to combine quantitative indicators with in-depth and well-informed narratives of each country's situation. The view here is that when domestic political conditions permit it and when such exercises have credibility, they can help promote progressive reform of policies and institutions.

## World Bank Country Policy and Institutional Assessments

The World Bank and the donor community have focused considerable attention recently on the CPIA as an index and as a process. The indicator has evolved over time as part of an ongoing effort by World Bank management to formally take account of borrowing country policy performance in allocating IDA resources. From the beginning, the Bank's intention was to develop a summary indicator that it could use to allocate greater amounts of IDA resources relative to need-based allocation criteria to countries that scored well on the indicator and relatively smaller amounts to poor performers. This was meant both to put resources into policy environments where they were expected to be relatively effective and as a way to encourage borrowing governments to improve their "performance".[5]

Driven by donors, the CPIA has mainly reflected their view on what constitutes appropriate policies and institutions for development. In practice, the CPIA reflects the views of the Bank staff members who make the individual country assessments and assign the scores.[6] The same staff members who are responsible for the Bank's programs in each country and the policy reform conditionality attached to those programs also make the CPIA assessments. This "by definition indicates the normative judgment of the World Bank as to which policy environments [are] best-suited to development" (Collier and Hoeffler, 2002, p. 26).

On the one hand, the CPIA is thus vulnerable to the civil society critique that it "rates governments on how faithfully they adopt neoclassical policies" (Alexander and Kessler, 2003). Indeed, a perusal of instructions to staff on how to grade countries gives much ammunition to this view (see World Bank, 2003). On the other hand, a number of the items that the CPIA seeks to capture should also appear on the list of essential policy and institutional concerns of the most heterodox advisors. Thus, even analysts who are critical of the full package of World Bank prescriptions that are summarized by the CPIA might find useful components that the exercise is seeking to measure.

In practice, however, and in a notable and continuing departure from the transparency that the Bank strongly advocates to its borrowing member governments, none of the CPIA country information is released to the public. Only World Bank staff members — not even executive directors — have access to the full CPIA. However, it will be argued here that there is reason enough in the weaknesses of the CPIA methodology for development analysts not to complain too loudly about this. Unfortunately, this seems hardly a satisfying answer as far as principles of governance of the World Bank are concerned.

## Development of the CPIA methodology

The aim of the Bank is that the CPIA assess "how conducive [a country's policy and institutional] framework is to fostering poverty reduction, sustainable growth and the effective use of development assistance" (World Bank, 2003, p. 1). This is to be accomplished by averaging scores on 20 aspects of a country's policies and institutions, the result being the CPIA index. In addition, the Bank clusters the 20 items into four sub-groups that are meant to summarize distinct categories of essential policy for development (see table 1). In essence, these clusters represent short-term economic management, long-term economic management, anti-poverty policy, and overall

[5] The quantitative ratings were never applied mechanically, but served to guide lending and to help management defend against pressures from individual borrowing countries and their bilateral supporters to increase allocations. In practice, the normative allocations were not always matched by programme outlays, and exceptions to the general allocation methodology were regularly applied, for example to limit allocations to the largest countries that might otherwise have absorbed most of what IDA had available to lend (see Kapur, Lewis and Webb, 1997, vol. I, pp. 1151-57).

[6] It appears from World Bank documents that the CPIA exercise includes middle-income countries that only borrow from the regular loan window of the Bank, as well as the "IDA-only" and "blend" countries.

domestic governance.

---

Insert table 1

---

Both the four clusters and the 20 individual items in table 1 seem to reflect less an overall coherent design than the history over a quarter century of step-by-step revision and accretion of concepts that management sought to include in the CPIA. Over time, individual items in the CPIA have been added and subtracted, split and merged. The main constant seems to have been that there be 20 items, and that they be weighted equally in the CPIA average. Thus, as new items came into the CPIA index, other items had to be collapsed or dropped.

The entry and exit of items has reflected both revisions in thinking about what are the most important elements of policy and changing pressures on management from IDA donors. One item in particular, "IDA portfolio performance" had an especially checkered career and perhaps is indicative of some of the pitfalls that can arise in designing quantitative performance indicators. This element entered the CPIA in 1993, as a result of a request by IDA donors to measure how well countries utilized specifically their IDA resources. According to the Bank's independent Operations Evaluation Department (OED), the only measure of portfolio performance available at the time was in reports of IDA supervisory missions, and this entered into the CPIA index with a weight of 20 percent. However, OED regarded the portfolio performance measure as "often subjective and biased in a positive direction", compared to ex-post OED evaluations (World Bank, 2001, p. 30). OED also regarded the IDA evaluations as not adequately separating the World Bank's own from the borrower's shortcomings. Moreover, if a badly performing project was dropped from a country's portfolio, the country's performance rating rose. In light of the above concerns, the weight of the portfolio performance indicator was reduced to 10 percent in 1995 and 7 percent in 1997, and it was dropped from the CPIA in 1998.[7]

It should be emphasized that from the start, the Bank sought to gauge policy measures taken and not development outcomes, which are not fully within the control of governments in developing or any other countries. The OED doubted that management succeeded in this regard and issued a warning against interpreting any internal Bank research as finding that "good policies" as measured by the CPIA from 1977 to 2000 help explain good economic growth (World Bank, 2001, pp. 16-18).

Perhaps the warning should be circulated (or re-circulated) to World Bank researchers, as they continue to use the CPIA as an explanatory variable in econometric exercises. For example, in a recently posted paper, Kraay and Nehru (2003) use the CPIA from 2001 back to 1977 — indeed, extrapolated back to 1970, based on an association of the CPIA with the domestic inflation rate — and claim that they have found a significant

[7] However, it still features in the performance-based allocation of IDA resources, as will be described in the annex to this paper.

inverse association of the quality of policies and institutions with the probability of debt distress. Perhaps they have and perhaps they found an association between high inflation and debt distress, given that inflation and the CPIA have been highly correlated, as noted in their extrapolation exercise, and given the OED observation on the tendency of Bank staff to rate countries with good economic outcomes as having good CPIA scores.[8] In sum, while institutions undoubtedly matter as a determinant of vulnerability to debt distress, econometric results using the CPIA do not necessarily show it.

Moreover, although the index was substantially revised in 1998 (and again in 2001) and smaller revisions are made each year, neither the changes in the structure of the CPIA nor in the definitions of individual items seemed to cause significant changes in the rating scores, at least through 2000. Indeed, the OED found the scores to be remarkably constant, even though there is a general impression that developing country policy has improved on average from the 1970s to today (World Bank, 2001, pp. 13 and 18).

The OED also found an unexpected shrinkage over time in the dispersion of CPIA scores, which it hypothesized resulted not from real policy movement toward the mean in the developing world, but from how the staff was coping with the increasing complexity of the CPIA scoring (World Bank, 2001, p. 18).[9] In addition, OED made another important observation about the CPIA that exemplifies a general caveat in the design of index numbers. That is, if each item in the CPIA is intended to have the same weight in the overall index, it should be normalized to have the same mean. In practice, however, some items tended to have higher scores and thus systematically had higher average weight in the CPIA than items with generally lower scores (World Bank, 2001, p. 20).[10]

**The CPIA methodology in 2003**

In light of these and other concerns, the Bank has devoted considerable additional staff time and resources to trying to strengthen the CPIA, with particular emphasis recently in developing indicators of the institutional dimensions that it now includes in the index. The results thus far, however, seem instead to confound different issues and sometimes challenge understanding.

As can be seen in table 1, some of the 20 items in the CPIA are policy indicators

[8] To be fair, the authors also experiment with a separate institutional quality measure prepared by the Kaufmann team, of which Kraay is part (see below), although it was only available for 2002, which gives them significant but weaker results than with the CPIA time series.

[9] One comment of OED in this context, which this author has seen repeated in other World Bank papers and is presumably the case, is that the staff making the CPIA assessments in the 1990s saw the exercise "as a means of establishing a rank order among countries rather than absolute scores" (World Bank, 2001, p. 18). This is curious because the country staff did not send ordinal rankings to CPIA management but cardinal scores. Also, presumably they knew that the ultimate use of the scores was in allocations of dollars, not to first place, second place, third place, etc. finishers, but to every IDA-eligible country, and for that an allocation based on ordinal rankings for some 80 countries would be incredibly cumbersome.

[10] By the same token, the items with higher variance of their scores have greater influence in the distribution of overall CPIA scores, which may be important for researchers with access to CPIA data to consider.

and others focus on institutions. However, not only are the two dimensions mixed in calculating each country's overall CPIA index number, but individual clusters also contain both institutional and policy dimensions. Cluster A, in particular, contains one fully institutional item (number 4), two essentially policy-related items (numbers 1 and 2) and one that is an equal mixture of both (number 3). In the latter case, the Bank instructs its country staff in assigning a score for the item to consider both "debt service capacity" and "debt management capacity" and weight each equally. Debt-service capacity pertains to the financial capacity to make timely payments to creditors, which depends on export earnings, capital flows, etc. and the policies regarding these variables that the government is pursuing. Debt management capacity pertains to having a central office to track and manage the government's financial obligations (see World Bank, 2003, p. 5). These really are quite different and it is not clear why they are merged rather than scored separately, other than the seemingly artificial commitment to keep the number of items averaged in the overall index at 20.

Indeed, a case could be made not only to separate the policy and institutional elements of individual items in the CPIA, but also to split the CPIA itself into two separate measures, one for institutions and another for policies. Having the institutional capacity to assess options, undertake policies and deliver public services effectively and fairly is quite distinct from policy choices that a government actually makes. Countries with strong institutions may elect governments that make unfortunate policy decisions. Equally, governments of countries with weak institutions may make appropriate policy decisions, albeit without necessarily being able to follow through fully and effectively on their implementation. It was the latter concern that led to the greater emphasis on institutional features in what had originally been a policy indicator, but they are distinct. Thus, while the Bank reasonably wants to track both institutional capacity and policy choice in its continuing country assessments, the CPIA as now constructed makes it very difficult to disentangle them.

The scoring of each of the 20 items is on a scale of 1 ("unsatisfactory for an extended period") to 6 ("good for an extended period"), with obvious gradations in between.[11] Staff members who make the assessments are given narrative guidelines that characterize the situations that should merit scores of 2 to 5 for each item. The staff are also given benchmark scores based on a subgroup of countries (19 in 2003), which are prepared as a preliminary round or pre-test of the CPIA assessment. To give the flavor of the guidelines, table 2 reproduces the full scoring guidance for two items, "Management and sustainability of the development program" (item 4) and "Property rights and rule-based governance" (item 16).

---

Insert table 2

---

[11] Possible scores are 1, 2, 2.5, 3, 3.5, 4, 4.5, 5 and 6, with 2 being "unsatisfactory", 3 being "moderately unsatisfactory", and equivalently for 4 and 5. The written instructions do not explain when or how to use the half-point scores. The "extended period" in scores of 1 and 6 is defined as at least 3 years.

The written instructions for the staff also include hot links to background research related to each characteristic, which the staff may consult for additional guidance. These include "objective indicators" that can serve as "guideposts" for making the subjective assessments. For example, the guideposts for item 8, "Competitive environment for the private sector", include such items as number of days needed and cost to register a business (World Bank, 2003, p. 11). In other cases, the guideposts are as judgmental as the CPIA item itself, or even more so. For item 17, "Quality of budgetary and financial management", the Bank staff member is directed not only to the IMF Code of Good Practices on Fiscal Transparency but also to a Checklist of Budget/Financial Management Practices that includes 23 items, each of which is to be assessed on a scale of 1 (inadequate) to 10 (excellent), including items like "Has medium-term perspective" and "Based on accounting standards" (World Bank, 2003, pp. 24 and 30). One must truly sympathize with the staff members of the Bank who have to make these assessments each year. The earlier observed regression to the mean in the scoring in the 1990s seems a fully human response to what management is asking them to do.

Moreover, while the 20 separate items in table 1 are clearly named, the actual content being measured is often not. For example, "financial stability" (item 6) seeks to capture three dimensions of policy believed to affect the degree to which a country is prone to financial crises. These dimensions are identified as competition policy, legal regime, and regulatory regime. The staff members are told in the instructions, furthermore, that one characteristic of competition policy that they should take into account is the degree to which the external capital account is open (World Bank, 2003, p. 9). But this seems odd. Most analysts would see that as a part of macroeconomic policy and neutral with respect to competition if applied appropriately (e.g., countries that wish to limit the degree of opening of their capital accounts, particularly regarding short-term flows, might still permit foreign-owned banks to operate in the domestic economy on the same basis as domestic banks).

One might also question if the dimensions identified in the instructions to staff for scoring a particular item are always the most relevant ones for that item. For example, if being more or less prone to financial crisis is the element that "financial stability" is meant to capture, then the first two of the three dimensions noted above seem out of place. Also, in addition to prudential regulations (the third dimension), it seems that ability of the central bank to be an effective "lender of last resort" would be significant. Indeed, the degree of openness of the capital account seems inversely related to the potential exposure to financial crises, especially for developing countries with thin financial markets.

By the same token, the competition and legal regime dimensions in item 6, as just cited, seem to fit better as dimensions of item 7, "financial sector depth, efficiency and resource mobilization." The focus there, however, is on monetary and credit policies, tax policies and ownership policies. Regarding the first set of policies, there is one focus that one expects to see in such instructions, namely to take account of the degree to which domestic interest rates are market determined. That is standard "Washington consensus". But it seems curious to also include the extent to which "the public sector borrowing requirement crowds out credit to the private sector" (World Bank, 2003, p. 10). The latter is a key macroeconomic issue that is already covered by item 1. In fact, maybe it is also covered in item 2.

In other words, not only is it not clear that each of the 20 items actually measures what it sets out to measure, but it is not clear that there are 20 distinct items. Item 1 "assesses whether a country has a consistent macroeconomic program (in terms of exchange rate, monetary and fiscal policy) that addresses inflation and internal and external imbalances." Item 2 "assesses the size of the fiscal balance and the composition of government revenue and spending to assess their compatibility with adequate provision of public services for economic growth, favourable macroeconomic outcomes, and a sustainable path of public debt" (World Bank, 2003, pp. 3 and 4). While there are aspects of item 1 that are not in item 2 and *vice versa*, the macroeconomic and debt sustainability aspects of item 2 seem part and parcel of the fiscal aspects of item 1. Much the same kind of overlap can be seen in the final two items (numbers 19 and 20) and for various items in between. However, if there are not 20 separate indicators in the CPIA, the index is not a simple average of 20 indicators but a weighted average of a smaller number of indicators in which the weights are the number of items that track a single factor.[12]

Whatever the assessment difficulties, scores are given for each item and are sent to the headquarters CPIA team, which checks for coherence and consistency in the scoring. This is important, as different staff members might well assign different scores to any one item for a country and a common standard would not necessarily result automatically. University teachers and students know this type of problem full well. Thus, a central team at Bank headquarters checks the scoring across countries of each individual item and reviews the narratives for assigning each specific score that now have to be sent to headquarters with the scores. The team also undertakes statistical tests for systematic bias in the scoring and examines outliers, assessing whether scores for such countries are warranted, possibly sending them back to country offices for revision.

**Use of the CPIA scores**

When the staff agrees to a final set of item scores for each country, the CPIA index numbers are calculated. Since 2000, World Bank country managers are required to share the 20 item ratings and overall index with their respective developing country counterparts. Each country's results are compared in a table to regional averages and to

[12] There are more and less precise statistical ways to ask how many separate items are really in the CPIA. At the more informal end of the spectrum, the OED reported finding substantial correlations, especially notable between the new governance items added in 1998 and the standard policy items (World Bank, 2001, p.20). In addition, in an early draft of an important paper arguing that aid should be targeted more on countries with good policies and high poverty, "good" was variously defined as relatively high scores on CPIA, cluster averages, and individual items in the CPIA. Each cluster alone "worked" in the growth and aid-effectiveness equation, as did 10 of the individual items, which the authors took to mean that the 10 contained almost all the information in the 20 (Collier and Dollar, 1998, Appendix). This was dropped in the published version of the paper and confidence in CPIA reliability as an indicator was placed instead on its having significance in a growth regression with common quantitative proxies for "good" policies, which were not significant (Collier and Dollar, 2002, p. 1499). This is a faith-based statement and not statistics, given the correlations among "independent" variables in the exercise.

their own performance in the previous year (when the instructions may have differed, at least for some items). The narratives that the assessing staff produced are also shared, providing a written commentary to explain the staff's thinking in assigning the scores. Developing country officials cannot challenge the scores. The only imaginable conversation is how to score higher next year, as the continuing central use of the CPIA is in allocating IDA resources (see annex).

No one else outside the Bank staff sees the scores or the narratives. Instead, beginning also in 2000, the Bank now groups the IDA-eligible countries into quintiles by their CPIA scores and reports the average score for each quintile (see table 3 for results of the 2002 exercise). Quintiles have similarly been calculated and reported for the four clusters of items. Observers can thus know if the World Bank judges a country to have relatively strong policies and institutions (defined as in the top fifth), relatively weak ones (bottom fifth) or something in between, and how "good" or "bad" each quintile is on average, based on the quintile mean.[13] One would not know from the data in table 3, however, that there were countries with ratings well below the mean of the lowest quintile or well above the mean of the highest quintile, nor that the distribution of CPIA scores is still very much bunched in the middle.

Insert table 3

The quintile report is a rather broad display of the results and one that World Bank management would like to improve upon. Indeed, as part of the most recent replenishment of IDA resources, the Bank is pledged to work to increase CPIA disclosure. As an intermediate step to eventual full disclosure, the Bank proposed in 2003 that it report country scores in half-point ranges (number of countries with scores from 3.0 to 3.4, 3.5 to 3.9, etc). However, when the IDA Deputies met in November 2003 and considered the matter of greater public disclosure, which they had requested, they were only able to agree to look "to the [Executive] Board for further discussion and decisions in advance of IDA-14, taking account of the diversity of views on the subject" (IDA Deputies, 2003).

The OED review of the CPIA asked already in 2001, why the controversy over full disclosure? Its answer was that "CPIA data are not yet robust enough to withstand full disclosure" (World Bank, 2001, p. 46). At that time, Bank staff did not have to write the narratives to explain their ratings and the OED thought that such an "audit trail" would greatly enhance the credibility of the ratings. Perhaps had the IDA Deputies seen these narratives, they would have held less diverse views in November 2003, or perhaps not.

[13] In an effort to emphasize that this is a grading of countries, one non-governmental organization translates the quintiles into letter grades (A,B,C,D,F) and publishes the result on the Internet (Citizens Network on Essential Services, 2003). The letter grades, however, go beyond what the data contain, as they assume the Bank is grading on a C curve and that the tails of the quintile distribution can be interpreted as "excellent" and "failing."

The question is whether the narratives would dispel the apparent concern for impartiality that is inextricably tied to the evaluations being done by Bank staff, however much those assessments are checked for consistency with assessments made of other countries by different Bank staff. The assessing staff member usually comes to the exercise with all the experience of having worked closely with the country. The staff member is thus very well informed from the Bank's side of the relationship with the government, which is necessarily a partial perspective. Moreover, the Bank staff members are in an asymmetrical power relationship with the government being assessed. This is not meant to impugn the integrity or competence of the staff in any way. It is about the ineluctable nature of the relationship.

## Other approaches to "enabling environment" indicators

It should be recalled that the initial and principal use of the CPIA was not analysis but allocating IDA resources, for which purpose, the views of the "lending officers" of the Bank were a naturally germane and confidential source (quantifying them into the CPIA is another matter). That origin also seems to suggest an important hypothesis as to why the index grew over time to be so complicated and difficult to interpret. In other words, at first, perhaps, some donors or management wanted to take account of the "quality" of macroeconomic policy in allocating IDA funds, and at a later time they or others may have wanted to add gender issues, and then property rights, and so on. In specifying the issues to include and agreeing to weight them equally, and even keeping a lid on the exercise by limiting the number of separate issues to 20, the concerns of various IDA donors and management could have been included in the allocation decisions in a practical way. If this is the process by which the CPIA was developed (and that history is not in the public domain), it would not be surprising if the average of the 20 items after 25 years did not make for an elegant or even coherent whole. But it also did not matter. Its purpose was resource allocation and not analysis. It had to produce an IDA allocation that seemed broadly correct to the donors and this it apparently did. At least, the IDA donors appear to want to retain it.

Other enabling environment indices seem to have either a political aim to influence policy or an informational one to assist global investors or for research purposes. In all cases, these indices are for public (or client) consumption. This allows them — actually, requires them — to be transparent about their methodology and reveal their calculations, as well as their detailed results. Some of them aim for a very broad coverage of issues and, like CPIA, appear to be attempting to characterize an "enabling environment" for development in a single index number. They can be contrasted with a more modest approach. The Global Competitiveness indices are examples of the first type and the governance measures of Daniel Kaufmann and colleagues at the World Bank exemplify the latter.

### Global competitiveness indices of World Economic Forum

The World Economic Forum (WEF) has developed two indices that are meant to compare national capacities for economic growth, based on a reading of the growth and development literature, in the case of one index and the business literature in the case of the second (WEF, 2004, chapter 1.1 and 1.2). The first index is based on sub-indices of the macroeconomic environment, the quality of public institutions and the development of technology. The second looks at the ability of enterprises to operate effectively (see table 4). These are complementary approaches (and the two indices are highly correlated), but they highlight different perspectives on requirements for development, reflecting the different professional disciplines of the authors of the two indices.[14]

---

Insert table 4

---

Together, the indices cover much the same area as the CPIA. The Business Competitiveness Index compares broadly with cluster B of the CPIA, and the Growth Competitiveness Index covers much the same ground as clusters A and D of the CPIA. The main difference is the explicit focus on technological advance and the absence of social indicators in the WEF, presumably reflecting its well-known corporate viewpoint. WEF is not a development bank.

Also, while the CPIA reflects the views of a small group of World Bank country experts, the WEF indices are based on both “hard” data (statistical time series, such as density of telephone lines as an indicator for ease of communication) and an Executive Opinion Survey. The latter is compiled by “partner institutes” in 102 countries (in 2003), following common guidelines for purposes of comparability. The partners are “typically leading national research or academic institutes committed to contributing to the growth potential of their respective economies” (WEF, 2004, p. xi). All together, the partner institutes surveyed almost 8,000 “senior business leaders” on the situations in their domestic economies in 2003. When, as for some low-income countries recently added to the exercise, the local business people surveyed did not appear able to adequately compare their economy to the rest of the world, a sub-sample of foreign business executives in that country was used instead.[15] All the survey information is necessarily subjective opinion and usually there is no alternative to it (e.g., extent of corruption would be hard to measure from published data).

There is considerable art as well as room for politics in constructing such indices. An example from the Growth Competitiveness Index illustrates the point. In previous years, the macroeconomic environment cluster contained an expression for government expenditure as a share of gross national product with a negative sign, meaning smaller was better. While this might be true at some high levels of government spending, it did not signal what was really important, distinguishing effective from wasteful spending. Recognizing this shortcoming, the authors this year instead sought to find a proxy for wasteful spending and ended up with a sub-index of the following survey questions

[14] Jeffrey Sachs and John McArthur developed the first index and Michael Porter the second. Xavier Sala-I-Martin and staff of the WEF produced the current version of the first and Porter continues on the second.
[15] The criterion for rejecting the local business views was when the standard deviation of their replies on a survey question significantly exceeded that for the global sample (WES, 2004, p. 36).

(WEF, 2004, p. 28):

- "Do government subsidies to business in your country keep uncompetitive industries alive artificially or do they improve the productivity of industries?
- "In your country, how common is diversion of public funds to companies, individuals or groups due to corruption?
- "How high is the public trust in the financial honesty of politicians?"

One can wonder if civil society advocates might have thought to ask respondents somewhat different questions, such as "how common is diversion of public funds … due to policy?" or "to what extent does the business viewpoint in general guide economic and social policy making?"

The WEF indices are also notable for rating all countries on a single scale, developed and developing. However, appearing on a single list is not the same as being on a single scale and the Growth Competitiveness Index also illustrates this well. In this case, the way technological progress was treated in the exercise undermined the technology cluster and the overall Index. Recognizing that technical change differs in advanced and developing countries, the authors divided up countries into more and less innovative economies, based on the number of US "utility patents" (for innovation) registered per capita in 2002. They then weighted the components of the technology index differently for the "core" (high innovation) and non-core countries, and also weighted the three major components differently in the overall index.[16] Thus, the position of individual countries in the overall index depends both on how they score on individual sub-clusters of items and how the clusters are weighted. Whether or not one finds them convincing, the two specific sets of weights make up separate indices and should be shown as such.

The Business Competitiveness Index, in contrast, is derived in a way that relies more on statistical methodology. First, each apparently relevant survey question and quantitative indicator for each sub-cluster in the index is tested for a significant correlation with GDP per capita. All those that passed this hurdle were then implicitly assigned weights in their cluster index through a factor analysis. Finally, the two main cluster indices were combined in a weighted average, where the weights (including an interaction term) resulted from a regression of the two indices on GDP per capita. While the author chose the initial questions to include in the exercise, he can say that the data chose how they would be represented in the final index.

In sum, the WEF indices give two business-oriented perspectives on an "enabling environment" for development. The surveys that underlie the indices are not of cross sections of local country populations, but of business leaders. This is an important segment of the population in every society, but not the exclusive one to sample, nor

[16] In the technology index, technology transfer has no weight in the core countries and the other items are weighted equally, while for the non-core countries, the innovation sub-index counts for 1/8, technology transfer for 3/8 and the weight of the rest is ½. In the overall index, the weight for the technology component is ½ for the core countries (the other components have ¼ each), while each of the three components receives equal weight for the non-core countries (WEF, 2004, pp. 5 and 27).

necessarily the one to rely upon, as Adam Smith implicitly warned us more than 200 years ago.[17] WEF then seeks wide dissemination of the business viewpoint captured in its indices to "help precipitate an internal debate within the country between government officials, business leaders, organizations of civil society and the academic community on key problem areas and how best to address them" (WEF, 2004, p. xi).[18] There is a political agenda in the WEF indices and it is a fully transparent one.

**The Kaufmann methodology: research driven**

The approach to constructing the WEF Business Competitiveness Index is part way to the more analytical research strategy followed by a group of World Bank researchers led by Daniel Kaufmann. They also generate an index — actually, indices — from subjective data, but they look for inputs from not one but 25 separate sets of data produced by 18 different organizations (Kaufmann, Kraay and Mastruzzi, 2003). Some of the data are surveys (including by the WEF) and others are expert assessments (including the CPIA). Some are global in scope and some have more restricted country coverage. All together, 250 variables pertaining to 199 countries and territories[19] were made candidates for indices on various hypothesized dimensions of "governance," which is the part of the "enabling environment" on which they focus.

As in the CPIA and WEF exercises, the authors begin with assertions of what the important dimensions of governance are, based on their "views of what constitutes a consistent and useful organization of the data that is concordant with prevailing notions of governance" (Kaufmann et al., 2003, p. 2). Thus, based mainly on the academic literature, they specify six indicators of governance: voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and control of corruption. These indices are not further combined into any overall concept of governance, nor do they feed into any overall "enabling environment" measure. Moreover, readers are warned repeatedly and even precisely about large margins of error in the results. Researchers in this team have their hubris under control.

The statistical procedure for combining the many data points into estimates of the six indicators of governance is itself interesting. The approach assumes that the six indicators are six true and unobserved variables and that the actual data points are linear functions of the true variable and a disturbance term. They then estimate the unobserved true variable from the observed variables and are even able to give greater weight in the estimation process to the observed variables that have stronger statistical properties (less "noisy signals"). In the Bayesian tradition, the authors also generate "confidence intervals" (with 90% probability) that the true and unobserved variable is within a specific range, as given by the probability distribution on the variable, conditional on the

---

[17] "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public…" (Smith, 1937, p. 128).

[18] Again, later in the report: "In the coming years, our aim is to work more and more closely with country leaders to improve the objectivity of the data collected in this *Report*, disseminate it more broadly, and create forums and other mechanisms to inform and catalyze local action" (WEF, 2004, p. 54).

[19] This was for the 2002 estimates; as some of the data were not available in earlier years, smaller data sets were used to generate estimates for exercises covering multiple years.

data used to generate it.

What is unique and highly important in this approach is that when the authors rank countries by their imputed scores from the estimated relationship, they not only can give the score itself, but also a confidence interval around it. Figure 1 shows this presentation for the “control of corruption” variable. The length of the confidence intervals vary, depending on how many data sources included that country and how much precision was ascribed to each source, but all the intervals are relatively large. The authors emphasize that small differences in the mid-point estimates between countries that are ranked near each other do not convey reliable information, although they are more confident in making comparisons of countries whose confidence intervals do not overlap.

---

Insert figure 1

---

Indeed, they suggest that users of their work “focus on the *range* of possible governance values” for each variable for each country, rather than the point estimate (Kaufmann et al., 2003, p. 13). They direct their concern in particular to efforts to allocate development assistance resources to countries on the basis of indicators such as these, citing in particular the announced allocation rules for the Millennium Challenge Account of the United States and the use of the CPIA in IDA allocations (ibid., p. 24).

## Conclusion

The starting point in this paper was the observation by many writers on development that the quality of institutions are important determinants of successful policy making and policy implementation for development. The next point was that somehow the quality of the institutions and of the policies could be measured. It seems that the main conclusion one might draw from the discussion above is that one should be very modest in what one claims regarding such measurements. One may get experts to quantify their opinions about a country or one may undertake an opinion survey of a more or less representative sample of people or one may even select an existing quantitative data series to be a good proxy for some aspect of institutional development.

None of that insures that the numbers generated mean what the designer intends them to mean or with an acceptable degree of reliability to warrant an analytical conclusion. The Kaufmann results contained an important warning about reliance on indicators such as the CPIA in allocating aid resources. Countries are just too susceptible to misclassification. The WEF effort reminds one to be sensitive to the policy agenda of entities promoting enabling environment indices. And from the CPIA exercise one sees the need for independent assessment and controlling the complexity of what one is trying to measure.

Certainly, one should be sensitive to the potential bias of experts making quantitative assessments,[20] or even of survey respondees (keeping in mind, for example, that the populations that the WEF samples do not reflect the full range of public opinion in a country). Another place where bias can enter is at the starting point of the exercise, when the analyst considers the broad institutional requirements against which the country is going to be compared. Do we really know what is required in an enabling domestic environment? Certain countries that do not compare well against success indicators like "rule of law" or "transparency" in government nevertheless seem to grow rapidly over extended periods, raise per capita real incomes substantially and notably reduce poverty. This suggests that perhaps there are multiple sets of effective policies and institutions for development. At the same time, one also has to be suspicious of too quick a retreat into cultural relativism. Political leaders of countries that lack certain characteristics have argued that view, which has rightly been dismissed, as by Amartya Sen (2003), as a dodge to deny their populations true global values, in particular ones that should be regarded as human rights.

Finally, what is left? The negative conclusions should not go so far as to write off the entire effort to quantify any aspect of institutional and policy development, but rather to appreciate their limits. There are other ways to approach the analysis. A traditional and useful one is to pre-establish a framework or set of questions within which to carry out the analysis of individual situations and then request independent experts to examine particular countries, utilizing the framework to structure their reports. Perhaps quantitative measures of aspects of the situation can usefully complement the structured narrative. Such a dual approach would allow the expert to be as nuanced as he or she wants in the narrative and to bring in features that would escape a quantitative approach, while the quantitative indicators can serve as a check on the analysis in selected topic areas that admit of such measurement.

In conclusion, our understanding of the package of necessary features of an enabling environment for development or whether there even are such things is at an early stage. The search for summary measurements seems to have gotten somewhat ahead of understanding. On the other hand, there are various components of an enabling environment on whose importance agreement can be reached and if measurement helps focus public debate on these items, so much the better.

[20] Kaufmann, Kray and Mastruzzi tested systematic bias of expert assessments from different institutions compared to results from sample surveys and found a consistent right-wing bias in Heritage Foundation assessments. However, in every instance in which a significant bias was found on any item, including in the CPIA on the issue of regulatory quality, it was in the right-wing direction. To be fair, the size of the effect was quite small. (Kaufmann et al., 2003, pp. 22-23).

**Annex. The performance-based system for allocating IDA resources**

Not only is the CPIA today a complex indicator that challenges interpretation, as was argued above, but the IDA allocation formula in which it is inserted also looks like a kind of "Rube Goldberg" invention. First of all, the CPIA does not even enter directly into the allocation formula but is first used to calculate a "Country Performance Rating" (CPR) and the square of the CPR is then entered into the formula for allocating IDA funds (see IDA, 2003, p. 2).[a]

There are several steps in calculating the CPR. The first is to combine the CPIA as the larger half of a weighted average of two indices. The other index seeks to track implementation of each country's active IDA projects and programs, and is produced as part of the Bank's overall Annual Report on Portfolio Performance (see World Bank, 2002).[b] In calculating the weighted average of the two indices, the CPIA is given a weight of 80 percent and portfolio performance receives 20 percent. The combined score is then multiplied by a "governance factor" which either raises or lowers the score, depending on whether the governance factor is greater or less than one. The result is the CPR.

The governance factor is calculated from the scores on CPIA items 4 and 16-20, plus a "procurement practices criterion" from the portfolio rating. The possible score on each of the seven items ranges from 1 to 6 and the first step is to calculate the simple average of the seven scores. This average is then divided by the average of the midpoint of the possible ratings of each item, i.e., 3.5. The resulting number is raised to the power of 1.5 and the result of that calculation multiplies the weighted average of the CPIA and portfolio performance. Thus, if a country's average score on the seven items lies above 3.5, the governance factor will raise its CPR above the weighted average of the CPIA and portfolio performance indicator. If the country's average score on the governance items is below 3.5, its CPR is lowered. Also, since the governance factor contains an exponent, the amount by which it raises or lowers a country's score increases exponentially, the greater the distance from the 3.5 "norm" (albeit only by a factor of 1.5). In sum, the expression for the CPR can be given as follows,

$$\text{CPR} = [.80\text{CPIA} + .20\text{PP}](\text{gov}/3.5)^{1.5},$$

where "gov" is the average of scores on the seven governance items and PP stands for the IDA portfolio performance measure.

---

[a] This author has seen descriptions of the formula, but not the formula itself. It is said to contain a basic fixed allotment for every IDA country, which helps increase the share of small economies, and a slowly decreasing function of gross national income per capita. In addition, "blend" countries get a lower allocation than the norm given by the standard formula in order to take account of their access, albeit limited, to other financial resources. Also, post-conflict countries and countries suffering from major natural disasters can receive larger allocations than given by the standard formula (IDA, 2003, p. 2).

[b] Recall that this was a component of the CPIA from 1993 to 1997. When it was dropped from the CPIA in 1998, it immediately reappeared as a separate factor that would be averaged with the CPIA in allocating IDA funds, as it remains today.

It may be seen that the governance elements of the CPIA can count rather heavily in the allocation exercise. In fact, the Bank acknowledges that as the governance elements enter into both the CPIA and portfolio performance ratings indices and again in the governance factor, they are double counted. Moreover, the exponential in the governance factor makes the IDA allocation quite sensitive to changes in the governance items. More precisely, the Bank calculates that a one point drop in one of the seven individual governance items will, *ceteris paribus*, reduce the Country Performance Rating by 7.5 percent and that will drop the country's IDA allocation by 15 percent (IDA, 2003a, p. 2).[c]

How the Bank's staff members rate the governance items thus matters. A relatively "easy grader" brings more resources to his country, albeit within the limits given by the CPIA headquarters consistency check. By the same token, the government will have a monetary incentive to undertake the governance reforms (or appear to undertake the reforms) that will score well in the CPIA. One may only wonder if there is a governance reform counterpart to the observation in United States school districts of "teaching to the test", so the average scores look as good as possible to the authorities and the voters. Unfortunately, that has not been the same as producing well-educated children.

## References

Alexander, Nancy (2004). "Country ownership undone." Citizens' Network on Essential Services, Silver Spring, Maryland. Unpublished manuscript (23 February).

Alexander, Nancy and Tim Kessler (2003). "The Millennium Development Goals (MDGs) in an unaccountable global order." December (on the web page of Citizens Network on Essential Services at www.challengeglobalization.org).

Calvo, Guillermo A. and Federic S. Mishkin (2003). "The mirage of exchange rate regimes for emerging market countries." *Journal of Economic Perspectives*, vol. 17, No. 4 (Fall), pp. 99-118.

Citizens Network on Essential Services (2003). "World Bank scorecard: tool to allocate resources among low-income countries in 2002." November (on the web page of the group at http://www.servicesforall.org).

Collier, Paul and David Dollar (2002). "Aid allocation and poverty reduction." *European Economic Review*. Vol. 46, No. 8 (September), pp. 1475-1500.

_____ (1998). "Aid allocation and poverty reduction." World Bank. October 20 draft (on the web page of the World Bank at http://www.worldbank.org/research/aid/background/bg98_allocation.htm and

[c] In fact, this result may have troubled Bank management, as it proposed to "slightly" reduce the impact of marginal changes in the governance ratings (IDA, 2003a, p. 2).

http://www.worldbank.org/research/aid/background/bg98_allocation_tables.pdf).

Collier, Paul and Anke Hoeffler (2002). "Aid, policy and peace: reducing the risks of civil conflict" (on the web page of the World Bank at http://econ.worldbank.org/programs/conflict/topic/12198/library/doc?id=12418).

Development Committee (2003). "Achieving the MDGs and related outcomes: a framework for monitoring policies and actions." Background Paper prepared by the staff of the World Bank and International Monetary Fund (DC2003-0003/Add.1), March 28.

Hout, Wil (2003). "Good governance and the political economy of selectivity." Working Paper No. 100. Asia Research Center, Murdoch University. October (on the web at http://wwwarc.murdoch.edu.au/wp/wp100.rtf).

International Development Association [IDA] (2003). "Allocating IDA funds based on performance: Fourth annual report on IDA's Country Assessment and Allocation Process." March (on the web page of the World Bank at http://siteresources.worldbank.org/IDA/Resources/PBAAR4.pdf).

IDA (2003a). "IDA's Performance-Based Allocation System: Current and emerging issues." October (on the web page of the World Bank at http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2003/10/27/000160016_20031027122520/Rendered/PDF/27082/pdf).

IDA Deputies (2003). "Chairman's summary." IDA-13 Mid-Term Review. Washington, D.C. November 5 (on the web page of the World Bank at http://siteresources.worldbank.org/IDA/Resources/IDA13MidTermReview.pdf).

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2003). "Governance matters III: Governance indicators for 1996-2002." Social Science Research Network Electronic Library, June 30 (on the SSRN web page at http://ssrn.com/abstract=405841).

Kapur, Devish, John P. Lewis and Richard Webb (1997). *The World Bank: Its First Half Century*. 2 volumes. Washington, DC: Brookings Institution Press.

Koeberle, Stefan (2003). "Should policy-based lending still involve conditionality?" *The World Bank Research Observer*, vol. 18, No. 2 (Fall), pp. 249-273.

Kraay, Aart and Vikram Nehru (2003). "When is external debt sustainable?" Preliminary manuscript. World Bank, October (on the web page of the World Bank at http://www.worldbank.org/hipc/progress-to-date/Debt_Sustainability/KraayNehruDebtSustainability10-16-03.pdf).

Sen, Amartya (2003). "Human rights and Asian values." Second Edition. Sixteenth Morgenthau Memorial Lecture on Ethics and Foreign Policy, New York, 1 May 1997. Carnegie Council on Ethics and International Affairs, New York.

Smith, Adam (1937). *An Inquiry into the Nature and Causes of the Wealth of Nations*. New York: Modern Library Edition.

United Nations (2002). *Financing for Development: Building on Monterrey*. New York: United Nations (Sales No. E.02.II.A.5).

World Bank (2003). "Country policy and institutional assessment 2003: assessment questionnaire," March (on the web page of the World Bank at http://siteresources.worldbank.org/IDA/Resources/CPIA2003.pdf).

World Bank (2002). "Annual Report on Portfolio Performance, Fiscal Year 2002." 2 volumes. December 12 (on the web page of the World Bank at http://siteresources.worldbank.org/QAG/Resources/arppfy02report.pdf; see also http://siteresources.worldbank.org/QAG/Resources/arppfy02Definitions.pdf and http://siteresources.worldbank.org/QAG/Resources/arppfy02statsappendix.xls).

World Bank (2001). "OED IDA Review: Review of the performance-based allocation system, IDA 10-12." 14 February (on the web page of the World Bank at http://wbln0018.worldbank.org/oed/oeddoclib.nsf/View+to+Link+WebPages/310529285EC4A83685256A680079BC9F/$FILE/PBA.pdf).

World Economic Forum [WEF] (2004). *Global Competitiveness Report 2003-2004*. New York: Oxford University Press.

**Table 1. Policy and institutional characteristics in the CPIA index, 2003**

A. Economic Management
1. Management of Inflation and Macroeconomic Imbalances
2. Fiscal Policy
3. Management of Public Debt (External and Domestic)
4. Management and Sustainability of the Development Program

B. Structural Policies
5. Trade Policy and Foreign Exchange Regime
6. Financial Stability
7. Financial Sector Depth, Efficiency and Resource Mobilization
8. Competitive Environment for the Private Sector
9. Goods and Factor Markets
10. Policies and Institutions for Environmental Sustainability

C. Policies for Social Inclusion/Equity
11. Gender
12. Equity of Public Resource Use
13. Building Human Resources
14. Social Protection and Labor
15. Monitoring and Analysis of Poverty Outcomes and Impacts

D. Public Sector Management and Institutions
16. Property Rights and Rule-based Governance
17. Quality of Budgetary and Financial Management
18. Efficiency of Revenue Mobilization
19. Quality of Public Administration
20. Transparency, Accountability and Corruption in the Public Sector

*Source*: World Bank, "Country policy and institutional assessment 2003: assessment questionnaire," March 2003, p. 1.

## Table 2. Scoring guidelines for two items in the CPIA, 2003

---

### Management and sustainability of the development programme (item 4)

| *Score* | *Guidance for score* |
|---|---|
| 2 | Institutions and instruments for implementing economic and development policies and managing shocks are not effective. Actual or incipient economic, political or security obstacles make it unlikely that authorities will implement needed reforms or maintain existing achievements. The public and key stakeholders have no influence on, or do not support, key decisions. The Government does not rely on participatory processes to gather information or review plans. |
| 3 | Institutions and instruments for implementing policies and managing shocks are weak, but there have been occasional successes. There are several impediments/obstacles that reduce the chances of successful reform efforts. There is only limited consultation and participation with key stakeholders and civil society. Participatory processes are rarely used, and public support is low. |
| 4 | Institutions and instruments work fairly well, although there are problems. While there are some impediments/obstacles to policy reforms, the Government has demonstrated the ability in the past to overcome these obstacles in many, but not all, cases. Some consultation has taken place with stakeholders, and participatory processes have been used in a limited fashion. Government coordination is good, and there is moderate public support for reform efforts. |
| 5 | Tools are available to implement policies and manage events effectively. Policies and actions of key agencies are well coordinated. Authorities have a coherent program of reform or a record of sustained good performance with broad public support. Participatory processes are often used as means through which the views of stakeholders can be heard and inform government decision making . |

### Property rights and rule-based governance (item 16)

| *Score* | *Guidance for score* |
|---|---|
| 2 | Enforcement of contracts and recognition of property rights depends almost entirely on informal mechanisms. Laws and regulations are applied selectively or changed unpredictably, for example through frequent and unpublicized executive decrees. Judicial decisions are not publicly available. Favoritism rather than equal treatment pervades dealings with the state. Obtaining a business license can take an inordinate time and require numerous "unofficial payments." Crime and violence substantially increases the cost of doing business. |
| 3 | The law protects property rights in theory, but in fact registries and other institutions required to make this protection effective function poorly, making the protection of private property uncertain. Judicial decisions are sometimes publicly available. Rules are not changed arbitrarily but may not be publicly available. Those without connections can secure a business license, but the process is overly bureaucratic and prone to delays. The state is able to provide a modicum of protection against crime and violence. |
| 4 | Property rights are protected in practice as well as theory. Property registries are reasonably current and noncorrupt. Rules are publicly available and a mechanism exists to resolve conflicts of rules. Courts may be costly to use but judicial decisions are publicly available. Obtaining |

necessary licenses is a small share of the cost of doing business. The state is able to protect most citizens most of the time from crime and violence.

5 A rule-based governance structure governs interactions between all citizens and their government. The legal system is highly predictable. Laws and regulations affecting businesses and individuals are transparent and uniformly applied; changes in them are publicly announced and occur only after public hearings and deliberation. A well-functioning and accountable police force protects citizens from crime and violence.

---

*Source*: World Bank, "Country policy and institutional assessment 2003: assessment questionnaire," March 2003, pp. 6 and 23.

**Table 3. Country Policy and Institutional Assessment (CPIA) ratings, 2002[a]**

| | |
|---|---|
| **First Quintile** Average = 3.69 | Bhutan, Cape Verde, Grenada, Honduras, India, Maldives, Mauritania, Samoa, Senegal, Sri Lanka, St. Lucia, St. Vincent and the Grenadines, Tanzania, Uganda, Vietnam |
| **Second Quintile** Average = 3.48 | Albania, Armenia, Bangladesh, Benin, Bolivia, Bosnia and Herzegovina, Burkina Faso, Ghana, Indonesia, Mali, Nepal, Nicaragua, Pakistan, Rwanda, Zambia |
| **Third Quintile** Average = 3.28 | Azerbaijan, Cote d'Ivoire, Dominica, Eritrea, Ethiopia, Kenya, Kyrgyz Republic, Lesotho, Madagascar, Malawi, Moldova, Mongolia, Mozambique, Republic of Yemen, Serbia and Montenegro |
| **Fourth Quintile** Average = 3.06 | Cambodia, Cameroon, Chad, Republic of Congo, Djibouti, The Gambia, Georgia, Guinea, Guyana, Kiribati, Niger, Papua New Guinea, Sierra Leone, Tonga, Vanuatu |
| **Fifth Quintile** Average = 2.57 | Angola, Burundi, Central African Republic, Comoros, Democratic Republic of Congo, Guinea-Bissau, Haiti, Lao PDR, Nigeria, Sao Tome and Principe, Solomon Islands, Sudan, Tajikistan, Togo, Uzbekistan, Zimbabwe |

*Source:* International Development Association, "Allocating IDA funds based on performance: Fourth annual report on IDA's Country Assessment and Allocation Process," March, 2003, p. 4.

[a] Countries are listed alphabetically within each quintile. IDA countries not rated in the 2002 CPIA exercise were Afghanistan, Liberia, Myanmar, Somalia, and Timor-Leste.

**Table 4. Policy and institutional factors in Global Competitiveness indices, 2004**

**Growth Competitiveness Index**

A. Macroeconomic environment
    1. Macroeconomic stability
    2. Government waste
    3. Country credit rating

B. Public institutions
    1. Contracts and law
    2. Corruption

C. Technology
    1. Innovation
    2. Information and communication technology
    3. Technology transfer

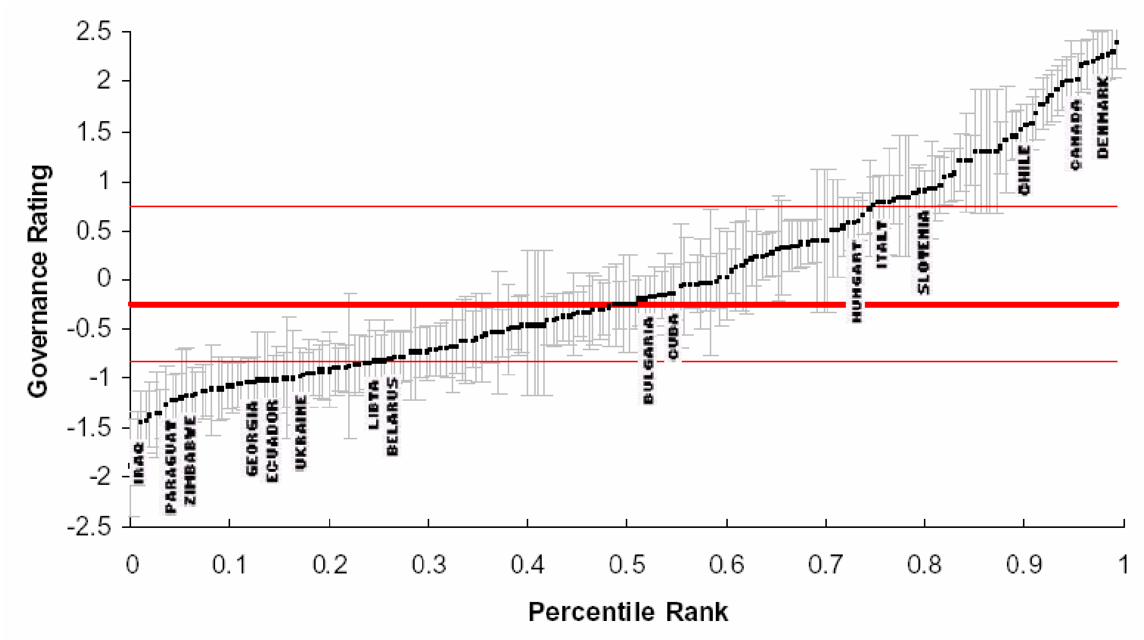
**Business Competitiveness Index**

A. Company operations and strategy

B. Quality of the national business environment
    1. Factor (input) conditions
        a. Physical infrastructure
        b. Administrative infrastructure
        c. Human resources
        d. Technology infrastructure
        e. Capital markets
    2. Demand conditions
    3. Related and supporting industries
    4. Context for firm strategy and rivalry
        a. Incentives
        b. Competition

*Source*: World Economic Forum, *Global Competitiveness Report 2003-2004* (New York: Oxford University Press, 2004), chapters 1.1 and 1.2.

**Figure 1. Indicator of control of corruption across countries, 2002**




*Source*: Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi, "Governance matters III: Governance indicators for 1996-2002," Social Science Research Network Electronic Library, June 30, 2003, p. 55.

*Note from the source*: "This graph shows estimates of the indicated dimension of governance (on the vertical axis) for all countries graphed against each country's percentile rank (on the horizontal axis) for 2002. The vertical bars show the statistically-likely range of values of governance for each country [90% confidence interval], with the midpoint of each bar corresponding to the best single estimate. Selected countries are labeled. As emphasized in the text, the ranking of countries along the horizontal axis is subject to significant margins of error, and this ordering in no way reflects the official view of the World Bank, its Executive Directors, or the countries they represent."