

Discussion Paper 3

Measuring Rights and Governance: The Quantification of Aid Conditionality

By Kate Desormeau

I. Introduction

In recent years, alongside the traditional watchdogs of human rights enforcement—international institutions like the UN, and CSOs like Amnesty International and Human Rights Watch—some of the main drivers behind increased global attention to human rights monitoring have been unlikely suspects: international financial institutions (IFIs) and wealthy donor governments.

Their interest is not academic. Official development assistance (ODA) worldwide suffered a steady decline throughout the second half of the 20th century, and it has only recently begun to rebound.¹ As official aid flows become an ever more highly public centerpiece of the global poverty reduction agenda, donors have a strong incentive to make the best possible use of the aid they are committing to developing countries—or at the very least to prove to their electorates that they are attempting to do so. The prevailing belief among donors and IFIs is that states exhibiting good governance and human rights practices will be more fertile ground for aid, yielding larger gains in terms of development, poverty reduction, and economic growth than more repressive or corrupt regimes. By this logic, donors eager to ensure the maximum impact of the aid they give developing countries should target it toward countries with good track records on human rights.

Thus, wealthy donor governments and IFIs are using human rights performance as one criterion for their aid allocations. This in itself is nothing new. Donors have long used human rights arguments to justify the giving or withholding of development assistance.² Until the late 1990s, however, donor governments relied mainly on subjective judgments that they applied at their discretion. Not surprisingly, this often had more to do with political and ideological motivations than with human rights principles for their own sake.³ As Katarina Tomaševski explained in 1997:

¹ United Nations Millennium Campaign, *Goal 8: More on Aid*, 2004, <<http://www.millenniumcampaign.org/site/pp.asp?c=grKVL2NLE&b=292089>>. See also OECD Development Assistance Committee, <http://www.oecd.org/linklist/0,2678,en_2649_33721_1797105_1_1_1_1,00.html>. Still, the UN observes, “when corrected for price and exchange-rate changes, the recent reversal in the decline in aid flows has barely brought real assistance back to the levels of 1990.” UN, *World Economic and Social Survey 2005: Financing for Development*, <<http://www.un.org/esa/policy/wess/>> at 110.

² This is also true of trade (for example, the application or revocation of “most favored nation” trading status, or the EU’s trade pacts with developing nations), but that is outside this paper’s scope.

³ One recent study, for example, analyzes Britain’s ODA patterns throughout the 1980s and 1990s to argue that human rights abuses only prompt donors to cut off aid when the “recipient is of little economic

Donor governments and agencies are continuously making decisions about which country to assist, how much and what for, because aid needs are much larger than available aid. Human rights have entered the already numerous criteria for allocating aid fairly recently. This entry has been neither easy nor smooth because no general criteria have been developed by donors and consequently decisions have been made on case-to-case basis. [. . .] Human rights criteria—to the extent that there is such a thing in the aid policy of any donor—have been confined to the search for those human rights violations which could justify cutting off aid.⁴

Recently, in contrast, there has been a push for such case-by-case conditionalities to be replaced by more systematic, “objective,” standards-based systems. In some cases, this push has manifested itself in the creation of statistical frameworks, composite indices, and ranking systems, which apply the same set of criteria to all would-be aid recipients in order to assess their performance in the areas of “human rights” and “governance.” Countries’ scores in these areas are then used to determine their eligibility for aid from year to year.

This push toward greater standardization is due in part to rising expectations of how efficient aid expenditures can and should be. Economists and development experts have produced a wealth of studies investigating the conditions under which aid is likely to have the greatest possible impact. Donors are eager commissioners and consumers of such studies, hoping to find more expedient ways of ensuring the maximum returns on their ODA investments. If this can be done through a single formula applied to all would-be aid recipients, then, from the perspective of some donors, so much the better. There is something undeniably attractive to policymakers about the idea of a universal formula for ODA success.

The push toward greater standardization is due, too, to the fact that ODA policy has become an extremely high-profile element of the international movement for poverty reduction. As citizens in donor countries become more aware of the importance of ODA in international politics and development—an awareness exemplified, for example, by the widely publicized Live Eight concert in July 2005—donors’ decisions have come under greater public scrutiny. It is now more important than ever before that donor governments be able to deflect allegations that their aid determinations are biased by political motivations. Thus, some donors have begun to bring statistical analysis to bear on their decisions, in the hopes that this greater standardization (and quantification) will make aid determinations appear more efficient and objective to their electorates, autonomous from political considerations.

value to the donor, and when the government of the donor state is politically weak” (43). See Bethany Barratt, “Aiding or Abetting: British Foreign Aid Decision and Recipient Country Human Rights” in Sabine C. Carey and Steven C. Poe, eds., *Understanding Human Rights Violations: New Systematic Studies* (Burlington: Ashgate 2004).

⁴ Katarina Tomaševski, *Between Sanctions and Elections: Aid Donors and their Human Rights Performance* (London: Pinter, 1997).

This paper will examine this new wave in human rights measurement. Looking principally at the measurement tools currently employed by the World Bank and the U.S. Millennium Challenge Account,⁵ we will consider where the discussion on ODA conditionality and human rights measurement is going, and how the human rights community must be involved in it. This paper argues that, in their eagerness to see returns on their investments in terms of poverty reduction, economic growth, and increased participation in the global economy, the World Bank and the Millennium Challenge Account have allowed measurement methodologies to run ahead of clear conceptual thinking about what exactly—human rights? governance? democracy?—they are measuring, and what data sources are best suited to capturing those phenomena. Reluctant to get involved in the ethical morass of aid conditionalities, the international human rights community has failed to help dispel this conceptual confusion. We have thus lost control of a critical part of our own mission. While we work to promote human rights around the world, what “human rights” means is being re-defined—without our input. This paper argues that the human rights community must add its unique voice to the debate on ODA, lest it find itself marginalized from the development agenda—and lest development be pursued in ways that fail to respect human rights.

III. The World Bank

For years, the World Bank has postulated a causal link between human freedoms and economic growth. In its *World Development Report* of 1991, the Bank demonstrated a positive correlation between “liberties” (as measured by Raymond Gastil’s human rights index, the precursor to Freedom House’s *Freedom in the World*) and “development” (represented by several proxies, including income growth, infant mortality, and the ratio of male-to-female educational attainment). While it drew no conclusions about the direction of causality, the Bank hypothesized that civil and political freedoms and development were interdependent.⁶

In the late 1990s, a range of new World Bank studies added support to this hypothesis, arguing that countries exhibiting “good policies”—including democratic governance and, in some cases, respect for civil liberties—were more likely to benefit from development aid.⁷ One study found that “rates of return on World Bank-financed

⁵ These two case studies have been chosen because they have the most comprehensive and “quantitative” systems for determining aid eligibility. Other countries and international organizations have also been developing measurement tools to help guide aid allocations, but space limitations do not permit us to cover them here.

⁶ World Bank, *World Development Report 1991: The Challenge of Development* (New York: Oxford University Press, 1991): 50, box 2.6. See also Russel Lawrence Barsh. “Measuring Human Rights: Problems of Methodology and Purpose,” *Human Rights Quarterly* 15:1 (February 1993) 87-121, and Gastil, “Comparative Study of Freedom.”

⁷ World Bank, *Assessing Aid: What Works, What Doesn’t and Why* (New York: Oxford University Press, 1998). Paul Collier and David Dollar, “Aid allocation and poverty reduction,” *Policy Research Working Paper* No. 2041 (Washington, D.C.: Development Research Group, World Bank, January 1999). Paul Collier and David Dollar, “Can the World Cut Poverty in Half? How Policy Reform and Effective Aid Can Meet the International Development Goals,” *World Development*, vol. 29, No. 11, 2001, pp. 1787-1802. Craig Burnside and David Dollar, “Aid, Policies and Growth,” *Policy Research Working Paper* No. 1777

projects in various developing countries over the 1974-93 period were higher in nations with greater civil liberties” than in countries with fewer civil liberties.⁸ The authors of this study concluded that their findings provided further evidence that “increasing public voice and accountability improves government performance” and the impact of ODA.⁹

Despite criticisms,¹⁰ this common-sense belief has prevailed in the Bank ever since: countries whose governments uphold “sound policies” (including the observance of good governance and civil and political rights) are likely to be more fertile ground for aid. They promise donors more bang for their buck in terms of poverty reduction and economic growth, and so they represent a wiser investment choice for the Bank’s development assistance. Thus justified, human rights and governance conditionalities make *economic*, not just *political*, sense. Their use therefore does not contravene the Bank’s Articles of Agreement, which bar the Bank from taking anything but economic factors into consideration when making policy decisions.¹¹ Following this logic, the Bank is free to use governance and human rights criteria to direct its aid-giving policies, mainstreaming these considerations into its aid eligibility computations. Below, we examine how it does so.

Computing eligibility

When the Bank allocates grants and loans to its borrowing countries each year, it takes into account three factors: the country’s CPIA rating (the Country Policy and Institutional Assessment, a calculation of the country’s adherence to 20 indicators of World Bank-approved fiscal policies),¹² its portfolio performance (how efficiently the

(Washington, D.C.: Development Research Group, World Bank, June 1997). Craig Burnside and David Dollar, “Aid, Policies and Growth,” *American Economic Review*, vol. 90, No. 4, September 2000, pp. 847-868.

⁸ The authors found no such correlation for “political freedoms,” however.

⁹ Jonathan Isham, Daniel Kaufmann and Lant Pritchett, “Civil Liberties, Democracy, and the Performance of Government Projects,” *World Bank Economic Review* 11 (May 1997): 219-42. See *Indicators of Civil Liberties*. 2001. The World Bank Group. 31 March 2005.

<http://www1.worldbank.org/publicsector/civil liberties.htm>.

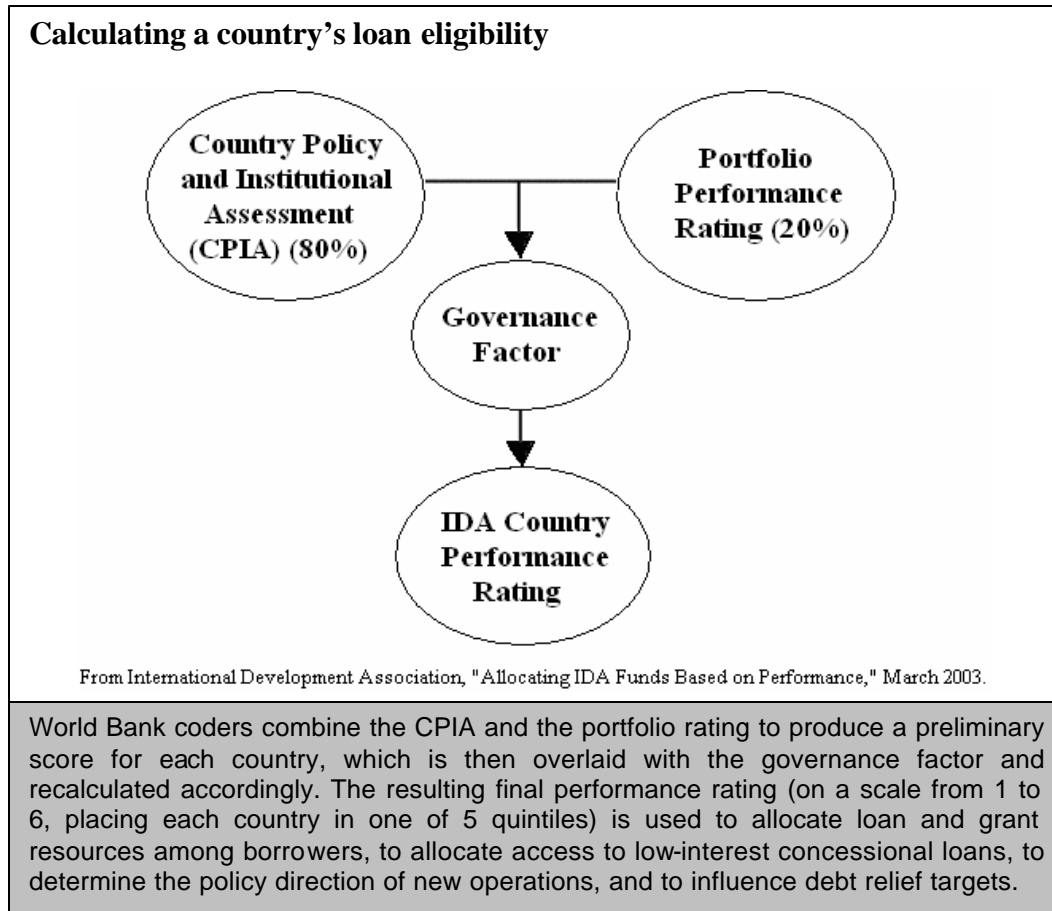
¹⁰ For a summary of studies attempting to debunk or at least problematize the Bank’s hypothesis, see UN, *World Economic and Social Survey: Financing for Development*, <http://www.un.org/esa/policy/wess/> at 119-20.

¹¹ World Bank Articles of Agreement, art. 4, sec. 10 (amended 1989).

¹² The CPIA scores countries on a scale from 1 to 6 on their adherence to World Bank standards (privatization, liberalization, etc.) in four clusters: (1) ECONOMIC MANAGEMENT (management of inflation and current account, fiscal policy, management of external debt, and management and sustainability of the development program); (2) STRUCTURAL POLICIES (trade policy and foreign exchange regime, financial stability and depth, banking sector efficiency and resource mobilization, competitive environment for the private sector, factor and product markets, and policies and institutions for environmental sustainability); (3) POLICIES FOR SOCIAL INCLUSION AND EQUITY (gender, equity of public resource use, building human resources, social protection and labor, and poverty monitoring and analysis); and (4) PUBLIC SECTOR MANAGEMENT AND INSTITUTIONS (property rights and rule-based governance, quality of budgetary and financial management, efficiency of revenue mobilization, efficiency of public expenditures, and transparency). See World Bank, “Country Policy and Institutional Assessment 2003: Assessment Questionnaire,” March 2003,

<http://siteresources.worldbank.org/IDA/Resources/CPIA2003.pdf> at 14. For a critique of the CPIA, see Barry Herman, “How Well Do Measurements of an Enabling Domestic Environment for Development

government has disbursed aid in the past to procure goods and services), and finally the so-called “governance factor.” The governance factor, being applied to the formula after the CPIA score has already been calculated and combined with the performance rating, is quite heavily weighted in comparison to the CPIA’s 20 constituent policy indicators, which are all averaged together.



The Kaufmann-Kraay governance factor

As shown above, once the Bank has determined a country’s CPIA rating and its portfolio performance rating, it applies a “governance factor” to arrive at the country’s overall rating for aid eligibility. The governance factor currently in use has been developed primarily by World Bank economists Daniel Kaufmann and Aart Kraay.¹³ The

Stand Up?” (presented at the XVIII Technical Group meeting of the Group of 24, Geneva, 8-9 March 2004) 3.)

¹³ This paper is based upon the 2003 version of the governance factor: see Daniel Kaufmann, Art Kraay, and Massimo Mastruzzi, “Governance Matters III: Governance Indicators for 1996-2002,” World Bank Institute Policy Research Working Paper 3106, 2003, <<http://www.worldbank.org/wbi/governance/pubs/govmatters3.html>>. Since this paper was written, however, the World Bank produced its next update of the governance indicators, “Governance Matters IV.” See <<http://www.worldbank.org/wbi/governance/pubs/govmatters4.html>>. For an overview of the changes included in this fourth edition, see the Center for Global Development, “Governance Matters: New Worldwide Indicators of Governance,” May 2005, <<http://www.cgdev.org/content/calendar/detail/3083/>>.

Kauffman-Kraay governance factor consists of a set of six aggregate indicator-sets that reflect six different dimensions of “good governance”:

- control of corruption,
- rule of law,
- voice and accountability,
- government effectiveness,
- political stability and absence of violence, and
- quality of regulation.¹⁴

For each of these six dimensions of governance, Bank coders consult data from a huge number of possible sources—including internal and external experts’ surveys and a range of qualitative and quantitative data from government agencies, CSOs, think tanks, and academia—to produce six aggregate scores for each country.¹⁵ The data sources used vary from country to country, depending on availability.

Of the six dimensions of governance, three—“voice and accountability,” “rule of law,” and “political stability”—appear to incorporate human rights considerations. These three dimensions rely on some explicitly rights-related data sources, including the annual reports of Reporters without Borders, Gallup International surveys, Freedom House country rankings, and finally a dataset the Bank calls simply “Human Rights.” This last dataset consists of information culled from the US State Department’s *Country Reports on Human Rights Practices* and Amnesty International’s *Annual Reports*, as organized and coded by the Cingranelli and Richards (CIRI) human rights index, a project run out of the State University of New York at Binghamton.¹⁶

Thus, when the World Bank is measuring “governance,” it is also interested in capturing certain aspects of human rights. But what are “rights” for the purposes of governance measurement, and how are they related to this broader governance phenomenon? Which rights-based indicators are essential to measuring “political stability,” for instance, and by what logic? Currently, the World Bank gives few clear reasons why it has selected the particular human rights sources it uses to assess a country’s score in this area. The Bank’s Articles of Agreement state that “only economic considerations shall be relevant” in its decision-making.¹⁷ The Bank only publicizes which data sources it uses for the governance factor, not which indicators it derives from them. Therefore, we do not know precisely which rights issues have been included and which have been omitted in the Bank’s final scores. All we can do is guess, by looking at the data sources used, which rights issues the Bank might be taking into consideration.

¹⁴ Daniel Kaufmann, Art Kraay, and Massimo Mastruzzi, “Governance Matters III: Governance Indicators for 1996-2002,” World Bank Institute Policy Research Working Paper 3106, 2003, <<http://www.worldbank.org/wbi/governance/pubs/govmatters3.html>>.

¹⁵ Country scores are an expected value of zero (standard deviation of 1; most values fall between -2.5 and 2.5). See The World Bank Institute, *Governance Research Indicator Country Snapshot (GRICS): 1996-2002*, 2005, <<http://info.worldbank.org/governance/kkz2002/>>.

¹⁶ Daniel Kaufmann, Art Kraay, and Massimo Mastruzzi, “Governance Matters III,” at 5.

¹⁷ World Bank Articles of Agreement, art. 4, sec. 10 (amended 1989).

And as we shall see in greater detail below, the preponderance of data sources used by the Bank deal with civil and political rights protection on a national scale. They do not appear to put a similar emphasis on group rights or on economic and social rights¹⁸, or to account for the distribution of resources and freedoms across a country (among various ethnic groups, geographical locations, religious groups, and so on).

The Bank's selective attention to some human rights and its omission of others is troubling from the perspective of a human rights practitioner or theorist, for it contravenes the idea of the indivisibility of rights—the idea that a state or duty-bearer cannot uphold certain rights and not others without ultimately compromising them all. Many human rights professionals and theorists would argue that rights cannot be imported one by one into a governance assessment; rather, “human rights” represents an interdependent framework of rights, the components of which are only meaningful and attainable when understood as a coherent whole. If human rights are to be part of the World Bank's concept of governance, does that mean *all* human rights (as enumerated in the UDHR, ICCPR, and ICESCR)? If not, which rights are essential, and which are not? Judging from the data sources currently in use, the Bank has bypassed this discussion, choosing instead to rely upon what data sources are available (and not what data sources contribute something essential to the concepts of human rights and governance). The Bank has thus allowed conceptual clarity to be held hostage to data availability. It uses data sources not because they capture essential aspects of rights—or because any sizeable body of research suggests that the particular phenomena they capture are indeed correlated with the effective use of ODA—but simply because they are available.

Moreover, the Bank does not disclose how it combines the various data sources it uses—all measuring different phenomena, all working under a different understanding of what “counts” in human rights—into a single “voice and accountability” or “rule of law” score. This approach raises the question of gaps and overlaps. Depending upon which data sources it relies upon, the Bank's calculations may miss certain essential elements of a rights issue while counting others multiple times. This will lead to a final score that is lopsided and foreshortened. If, for example, Freedom House and Reporters Without Borders both contain some measure of state-sponsored violence against independent journalists, but neither gives an indication of how literacy rates are distributed across the country's regions—which many human rights advocates would argue is just as essential to ensuring that all citizens can have a voice in their government—then is the resulting “voice and accountability” score disproportionately focused on only one relevant aspect of the rights issue in question while neglecting another? Outside the Bank, we have no way to find out how exactly these composite scores have been put together, what they may be missing and what they may be over-emphasizing.

No data source is inseparable from the interests, emphases, and agendas of its creators. The Kauffman-Kraay system assumes that the inevitable biases and selective emphases of its data sources will cancel one another out simply by virtue of the sheer

¹⁸ This is not to say that the World Bank indicators overlook economic and social *development* indicators, but that they do overlook economic and social *rights* indicators. The difference is discussed in Andrew Hines's paper, “Central Issues in Human Rights Measurement.”

number of sources used. But this scattershot approach to is unreliable; the Bank has not proved that merely adding *more* sources truly succeeds in eliminating their biases and yielding a balanced final score.

This unreliability is compounded by the fact that quite different numbers of data sources are available for different countries. This is perhaps inevitable given the wide variations in data availability worldwide, but it should make one wary about the comparability of the results. The margin of error for country A, for which there may be few acceptable data sources available either because it is not economically important enough to warrant international attention or because its government has actively discouraged the presence of international monitors working there, will be much larger than the margin of error for country B, on which a large number of CSOs, business monitors, and other “external experts” produce reports that can be cross-checked with one another. The potential for differences in the quality of data available for each country is worrisome, threatening to further marginalize countries to which international monitors have only limited interest or access by making them less eligible for aid.

A closer look: Indonesia’s “voice and accountability” score

Let us look more closely at the Bank’s choice of data sources in one governance dimension in particular: “voice and accountability.” According to Kaufmann and Kraay’s methodological report, this category is intended to measure

various aspects of the political process, civil liberties and political rights [and] the extent to which citizens of a country are able to participate in the selection of governments. We also include in this category indicators measuring the independence of the media, which serves an important role in [. . .] monitoring those in authority and holding them accountable for their actions.¹⁹

This description would seem to coincide quite neatly with a traditional conception of civil and political human rights. Interestingly, however, the Bank’s choice of data sources reveals a focus that most human rights professionals would find unfamiliar. The data sources used here mix mainstream human rights indicators with business, economic, and corruption indicators that are meant to capture regime and market stability and make projections about investment risks in foreign countries. Consider Indonesia, for example, whose “voice and accountability” score is computed based on 10 separate data sources:²⁰

Source	Author or sponsor	What it measures
State Capacity Project	Columbia University	Analyzes countries’ political context, state legitimacy, human resources and organizations, institutions, and overall capacity based on external experts’ polls.
Country Risk Service	Economist Intelligence Unit	Analyzes the business, economic and political risk factors based on external experts’ polls.

¹⁹ Daniel Kaufmann, Art Kraay, and Massimo Mastruzzi, “Governance Matters III,” at 4.

²⁰ See the World Bank Institute, *GRICS: Governance Research Indicator Country Snapshot*, <<http://info.worldbank.org/governance/kkz2004>>.

Freedom in the World report	Freedom House	Ranks countries' performance in civil and political rights based on external experts' polls.
Voice of the People survey	Gallup International	Surveys random population samples on their perceptions of democracy, human rights, values, crime, and the environment.
Global Competitiveness Report	World Economic Forum	Asks business executives to rate countries on a scale of 1 to 7 in openness, government, finance, infrastructure, technology, management, labor, and institutions.
Human rights reports	"Human rights" (State Department/Amnesty International/CIRI database)	Assesses countries' human rights practices qualitatively (using expert opinions, eyewitness accounts, etc.). Qualitative evaluations are then coded quantitatively.
International Country Risk Guide	Political Risk Services	Assigns scores of political, financial, and economic investment risk based on external assessments.
Press Freedom Ranking	Reporters without Borders	Assigns country rankings based on a questionnaire on press freedom sent to internal and external experts.
World Competitiveness Yearbook and Scoreboard	Institute for Management and Development	Ranks countries' business climates (in terms of economic performance, government efficiency, business efficiency, and infrastructure) based on external experts' polls and objective data.
World Markets Online	World Markets Research Center	Rates the quality of the business and investment climate in countries worldwide.

Clearly, the sorts of indicators used here to assess "voice and accountability" in Indonesia are not primarily rights-focused. Instead, certain civil and political rights are collected and taken into account as subsidiary components of a governance evaluation, while other rights (like domestic violence and ethnic discrimination) are not.²¹ Furthermore, all the data sources dealing with socio-economic conditions are not rights-based—that is, they are not designed to identify intent and accountability of duty-bearers, and they are not disaggregated on an intra-country level as rights-based indicators should be.²² Thus, judging from the Bank's choice of data sources, "governance" seems to be a concept that aims to capture not equity, but political and economic stability above all.

The Bank's selective attention to rights is not surprising, or even necessarily blameworthy. The Bank does not claim to be interested in "human rights" for their own sake, but rather for what they tell us about aid effectiveness and economic growth. But by what logic are the rights included in "voice and accountability" (like press freedom) relevant to predicting economic growth and the favorable impact of aid, while other

²¹ Of course, due to the Bank's non-transparent methodology, it is difficult to tell what exactly these data sources are being used for, and what indicators are extracted from them. Judging from the thematic emphases of Reporters without Borders, Freedom House, the US State Department, and to a lesser extent Amnesty International, we can safely surmise that the "human rights" the World Bank is interested in belong overwhelmingly to the civil and political sphere.

²² See Andrew Hines's working paper, "Central Issues in Human Rights Measurement," for a discussion of what makes indicators "rights-based."

human rights concerns (like minority access to voting rights) appear to be omitted? It is not clear that the rights issues represented in the voice and accountability score, and those alone, are in fact correlated with economic growth or poverty reduction. In the absence of conclusive research one way or another, the human rights community should ask itself: are we comfortable with the Bank's selective attention to the rights we consider to be indivisible?

Measuring “governance” or “human rights?”

As the example of Indonesia above demonstrates, governance and human rights are two distinct concepts, and measurement approaches differ accordingly.²³ Most of the data sources that are used to compute Indonesia's score are clearly not formulated with rights in mind. Rights-based indicators should be concerned with the claims of individuals and the corresponding responsibilities of duty-bearers. In contrast, governance and political risk indicators (such as the ones listed above) tend to be more concerned with national aggregates, and they may not pick up on disparities in the welfare of regional, ethnic, or religious minorities. Furthermore, because of their focus on law, capacity, and institutional frameworks, governance indicators are frequently more concerned with structure (laws and regulations passed, remedies available, etc.) and aggregate outcomes (national levels of attainment) than with process (actual government activities, improvement over time).²⁴ Thus, they fail to achieve the balance of structural, process, and outcome indicators that Paul Hunt has shown to be necessary for rights-based measurement.²⁵

Governance measures also envisage governments as the primary duty-bearers, and may overlook the responsibilities of other, non-state actors who can also affect the enjoyment of rights (like militia groups, corporations, and even donors themselves). Thus, governance indicators can have a sort of tunnel vision which might be appropriate to the task of identifying countries whose official policies are likely to foster economic stability and growth, but which might not be able to pick up on other things—like equity and justice, the enjoyment of rights of the disenfranchised and their actual chances of redress when these rights are violated—which are relevant to a true human rights analysis, and which indeed may be relevant when predicting whether ODA will have long-lasting benefits in receiving countries.

²³ Todd Landman, “Measuring Human Rights: Principle, Practice, and Policy,” *Human Rights Quarterly* 26 (2004): 929.

²⁴ Steven Radelet *et al.* have criticized the Millennium Challenge Account for this reason, pointing out that its emphasis on levels of attainment over rate of improvement may penalize countries that are taking steps in the right direction simply because the results of their efforts have not yet materialized. See Steve Radelet, Sarah Lucas, and Rikhil Bhavani “A Comment on the Millennium Challenge Account Selection Process,” Center for Global Development, 9 March 2004, <<http://cgdev.axion-it.net/docs/CGD%20Comment%20on%20MCA%20Selection%20Process.pdf>>, at 3.

²⁵ See Paul Hunt, *Interim report of the Special Rapporteur of the Commission on Human Rights on the right of everyone to enjoy the highest attainable standard of physical and mental health*, United Nations General Assembly, Fifty-eighth session, Agenda item 117 (c), 10 October 2003, United Nations General Assembly, paragraphs 14-29. See also Rajeev Malhotra and Nicholas Fasel, “Quantitative Human Rights Indicators: A Survey of Major Initiatives,” draft for discussion at Turku, 3 March 2005. <<http://www.abo.fi/institut/imr/indicators/Background.doc>> para. 63-5.

Given the divergent aims of human rights and governance measurement, it is worth questioning whether, in compiling “voice and accountability” scores, the World Bank is right to mix measures of political stability, investment risk, and economic growth potential on the one hand, and more traditional “human rights” measures of civil and political freedom on the other, without distinguishing between their differing aims. Freedom House, Amnesty, the US State Department, and Reporters without Borders all claim to be measuring aspects of human rights, not governance. Yet the Kauffman-Kraay dataset derives information from them and inserts it into a governance framework without providing a rationale for thus transporting data from one conceptual field to another. The database does not attempt to explain exactly how it combines these diverse concerns into a single composite score, or how these concerns should be weighed relative to one another. How many units of press freedom, for example, are equal to how many units of economic openness? Should press freedom, or economic openness, count for more of the overall score in countries where fewer data sources are available? As Landman and Häusermann have noted, the governance measurement endeavor is beset by such questions of conceptual confusion.²⁶

The Bank’s human rights data sources

The Bank’s reliance on the above data sources entails further problems, these ones related to characteristics of the data sources themselves. Certainly, of all the rights-based analyses available today, Amnesty International’s *Annual Reports* and the US State Department’s *Country Reports on Human Rights Practices* reports offer unusually broad geographical and temporal coverage. They are thus understandably attractive candidates for inclusion in a system of measurement like the Kauffman-Kraay, which wants to incorporate some aspects of human rights in its computations. This does not mean, however, that the sources are methodologically suited to the task for which they are being employed.

As Eleanor Benko shows,²⁷ Amnesty’s reports are constrained by a number of extrinsic considerations (i.e., space limitations and funding concerns) and by intrinsic, methodological considerations (i.e., the difficulty of conveying the relative severity of violations across countries or over time in a purely qualitative report). Intended principally as tools for research and advocacy, Amnesty’s *Annual Reports* are not meant to guide aid allocations. While the State Department’s human rights reports *were* initially commissioned for the purpose of guiding aid policy (see section IV below), they too are unable to convey a systematic sense of time-series. Like Amnesty reports, they assess each country’s progress or deterioration qualitatively and anecdotally, leaving the reader with no concrete sense of change over time.²⁸

²⁶ Todd Landman and Julie Häusermann, “Map Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance,” report for the Statistical Office of the Commission of the European Communities (EUROSTAT), University of Essex: Human Rights Centre, 2003, <<http://www.oecd.org/dataoecd/0/28/20755719.pdf>>.

²⁷ See Eleanor Benko’s working paper, “How Human Rights Organizations Measure Human Rights.”

²⁸ As Eleanor Benko’s paper points out, Amnesty reports cannot be read as time-series. Neither can State Department Reports, for similar reasons—although the State Department itself would probably dispute this,

For the purposes of the World Bank governance factor, however, these qualitative reports are coded as though they were quantitative and reliably sensitive to changes over time. Coding is done by the CIRI Human Rights Dataset, a project funded by the Bank and based at SUNY Binghamton. Coders analyze the text of State Department and Amnesty annual reports, assigning numerical scores to each country in a number of rights-related areas (physical integrity, civil liberties, workers' rights, and gender empowerment) based on the narrative descriptions in those reports.

The relative prevalence of extrajudicial killings from year to year and country to country, for example, is determined as follows. If the US State Department's report provides numbers on extrajudicial deaths, then a country where no such deaths occurred in a year would receive a score of 2 (no violation). Between 1 and 49 deaths would correspond to a score of 1 (occasional violations), and 50 or more deaths would correspond to a score of 0 (frequent violations). If the State Department report does not include any numbers for a particular rights issue, then the coders must rely on whatever descriptive words are used to gauge the prevalence of extrajudicial killings. The CIRI coding manual explains that adverbs and adjectives like "endemic," "gross," "extensive," and "routinely" correspond to a score of 0 (frequent violations). When confronted with words like "numerous," "many," "occasional," and "various," coders must "use [their] best judgment" to determine whether such a description deserves a score of 1 (occasional violations) or 0 (frequent violations).²⁹ The resulting country scores are then used to make comparisons across countries and over time.

It is these CIRI scores, among other datasets, that Kauffman and Kraay use to calculate the Bank's country scores for "voice and accountability," "political stability," and "rule of law." They are reassessed periodically and used to adjust aid allocation decisions from year to year. Thus, quite arbitrary and subjective factors such as word choice, thematic emphasis, or word count limits (which are part of any narrative reporting framework) may have real implications on how a country fares in its governance scores. Given these shortcomings—and the very practical, dollars-and-cents implications they have for low-income countries—the use of these data sources may be inappropriate.

In measuring "governance"—that is, assessing a country's institutional capacity to absorb and utilize aid—the Bank has not fully worked out how measures of civil and political, economic and social, and group rights protection fit in. Yet despite the conceptual and methodological problems outlined above, human rights measures—ready or not—have been subsumed into the World Bank's governance-measuring framework, and have become a central tool for assessing eligibility for aid. The effects of this practice reach even beyond the Bank's own policymaking. The Bank is a pace-setter for bilateral aid allocations. Donor countries often follow the Bank's (and one another's) lead when

since it does claim to show "progress" in countries over time. Similar questions could be raised about Reporters without Borders and Freedom House, although Freedom House does self-consciously keep its methodology consistent from year to year to reduce the problem of coding inconsistencies.

²⁹ The Cingranelli-Richards Human Rights Database Coder Manual, version 8.01.04, 8. http://ciri.binghamton.edu/web_version_7_31_04_ciri_coding_guide.pdf.

the Bank reduces or denies funds to a borrowing country. This tacit “herding” phenomenon means that each of the Bank’s decisions can have wide ripple effects, making some countries favorite aid recipients while others become “aid pariahs”—which, ironically, are often the countries whose populations are most in need of assistance.³⁰ Some bilateral donors explicitly incorporate World Bank assessments into their own eligibility computations. For example, the United States has adopted the Kaufmann-Kraay “Voice and Accountability” indicator cluster for use in directing its Millennium Challenge Account aid, as we shall now see.

IV. The United States

Official development assistance—channeled not only through the US Agency for International Development (USAID) but also through the State Department, the Department of Defense, and other government agencies—is a key tool of US diplomacy abroad. Acting on the belief that democratic countries are inherently more peaceful and stable,³¹ the US government targets development assistance toward countries exhibiting signs of “democracy,” “good governance,” and “human rights” (and denies aid to those who do not) in the interest of combating terrorism and fostering international security. These political motivations are supplemented, as always, by economic ones: US official development assistance as a percentage of GDP fell steadily throughout the second half of the 20th century, and only recently began to rebound.³² In the context of this economic history—and in line with the rationale that funding should be targeted toward the countries that will make the most efficient use of it—measures of human rights are frequently used to ration and rationalize the aid the US now gives.³³

³⁰ The Netherlands in 1999 explicitly began relying upon the CPIA’s quintile rankings to guide its own bilateral ODA allocations. See Barry Herman, “How Well Do Measurements of an Enabling Domestic Environment for Development Stand Up?” 3. See also Wil Hout (Institute of Social Studies, the Netherlands), “Good Governance and the Political Economy of Selectivity,” Murdoch University working paper number 100, Perth: Asia Research Centre, October 2003, <<http://www.bothends.org/politiekcafe/polcaf18-wp100.pdf>>.

³¹ USAID, *Foreign Aid in the National Interest: Promoting Freedom, Security, and Opportunity*, Washington DC, 2002, <<http://www.usaid.gov/fani/>>.

³² Between 1949 and 1995, US ODA as a percentage of GDP fell from a Marshall-Plan high of 3% to a low of 0.17%. Today, despite considerable improvement in absolute terms between 2000 and 2004, the percentage stands at just 0.16% (\$19 billion). In this, the United States lags behind several European countries, and of course falls far short of the UN’s goal of 0.7%. See USAID, *US Leads Rise in 2004 as Global Aid Donations Increase to \$78.6 Billion*, <http://www.usaid.gov/press/frontlines/fl_may05/donations.htm> and USAID, *Untitled*, 1995, <<http://www.usaid.gov/pubs/apr95/apr2.pdf>> and Ralph Dannheisser, USAID, *Head Lauds Increase in U.S. Aid to Africa, Promises More*, <<http://tokyo.usembassy.gov/e/tp-20050705-18.html>>. See also Rikhil Bhavnani, Nancy Birdsall and Isaac Shapiro, “Whither Development Assistance? An Analysis of the President’s 2005 Budget Request,” budget memo, Center for Global Development and Center on Budget and Policy Priorities, July 2004. <<http://www.cgdev.org/docs/DAbudgetnote05.pdf>>. See also “America, the Indifferent,” editorial, *New York Times*, 23 December 2004.

³³ USAID uses a large set of democracy and governance indicators to track the impact of its aid strategies in developing countries. See Center for Democracy and Governance, “Handbook of Democracy and Governance Program Indicators,” USAID Technical Publication Series Document PN-ACC-390, 1998. <<http://www.usaid.gov/democracy/pubsindex>>. The handbook includes about 350 indicators.

It was the Carter administration that firmly incorporated human rights as a decisive factor in US foreign aid policy. In the 1970s, in response to Congress's requests for country information on which to base trade decisions, the State Department began issuing its annual *Country Reports on Human Rights Practices*. These are qualitative, country-by-country assessments written by State Department desk officers, drawing primarily on information gathered from US embassies abroad. Under President Reagan, concern with "human rights" was partially eclipsed by concern with "democratization."³⁴ Whether the United States is promoting "human rights" or "democracy" abroad has been a bone of contention ever since. Predictably, some critics argue that the State Department reports are biased and politically motivated. Criticism to this effect has never been so heated as this year, 2005, when several foreign governments issued official statements suggesting that the administration that had presided over abuses at Guantánamo and Abu Ghraib was in no position to evaluate other governments' human rights practices.³⁵

Accusations of bias notwithstanding, the US government has authorized two more high-profile, annual human rights reports in the past ten years that closely mimic the State Department country reports. These are the *International Religious Freedom Report* and the *Trafficking in Persons Report*.³⁶ Both reports isolate one particular rights issue (religious freedom or trafficking), solicit information from US embassies and local sources, and use that information to issue qualitative, narrative accounts of the health of that right in countries throughout the world. These analyses can be employed by the US government to guide diplomatic relations generally, as well as to make ODA allocation decisions, to offer discretionary funding, and to call for sanctions.

The International Religious Freedom Report

In 1998, the International Religious Freedom Act established the Office of International Religious Freedom in the State Department. The Office is responsible for compiling an annual global report that describes the state of religious freedom in 191 countries.³⁷ On the basis of these narrative country reports, the Secretary of State designates "countries of particular concern" (CPCs) and notifies Congress of her decisions, which can then be used as the basis for economic sanctions.³⁸ While the report

³⁴ John Shattuck, "Diplomacy with a Cause: Human Rights in US Foreign Policy," in *Realizing Human Rights: Moving from Inspiration to Impact*, ed. Samantha Power and Graham Allison (New York: St. Martin's Press, 2000) 269.

³⁵ Edward Cody, "China, Others Criticize U.S. Report on Rights: Double Standard at State Dept. Alleged," *Washington Post Foreign Service*, 4 March 2005.

³⁶ The US Department of Labor has also begun issuing annual *International Child Labor Reports*. See <<http://www.dol.gov/ILAB/media/reports/iclp/main.htm>>

³⁷ See <<http://www.state.gov/g/drl/rls/irf/2004/>>. "The goal of this report is [. . .] to create a comprehensive record of the state of religious freedom around the world. [. . .] It is our hope that this report will [. . .] provide all three branches of the federal government as well as the press, foreign governments, religious groups and NGOs with a factual basis for evaluating religious freedom worldwide" (Harold Koh, then assistant secretary of state for democracy, human rights and labor, transcript *EPF411, 9 September 1999, <<http://canberra.usembassy.gov/hyper/WF990909/epf411.htm>>.

³⁸ Sanctions are considered to be a "last resort." The country reports are also used more generally to guide US diplomacy. The IRFA also created a U.S. Commission on International Religious Freedom

is generally considered reliable, some critics do argue that the CPC designation process is biased. For example, Uzbekistan—which enjoys a special relationship with the United States due to its geo-strategic importance in the war on terror—has never been named a CPC despite what most observers consider to be an atrocious religious freedom record.

The Trafficking in Persons Report

Established by the Trafficking Victims Protection Act (TVPA) of 2000, the Office to Monitor and Combat Trafficking in Persons (also housed within the State Department) issues an annual *Trafficking in Persons Report*.³⁹ Like the *International Religious Freedom Report*, the trafficking report is a country-by-country compilation of qualitative assessments. Due to the clandestine and highly controversial nature of the trafficking phenomenon and the difficulty of compiling reliable statistics, country narratives focus on process (laws passed, anti-trafficking policies in place) rather than on outcomes (estimating how many people are in fact trafficked into or out of the country). Based on these narratives, the report places each country in one of three tiers. Tier One countries are in full compliance with TVPA's minimum standards; Tier Two are not in compliance but are making reasonable efforts; and Tier Three are making insufficient efforts. Country assessments can be used to target the money that the TVPA makes available for anti-trafficking assistance programs overseas.⁴⁰ The President may declare a Tier Three country ineligible for these funds, and he may also use the Tier Three determination to pressure the IMF to withhold its non-humanitarian and non-trade related assistance.⁴¹ Critics observe that this is problematic. Naturally, it does no good to reward countries that refuse to implement anti-trafficking policies by giving them funds that they will not use in good faith. Nevertheless, since trafficking is considered to be strongly correlated with a lack of domestic economic opportunity,⁴² the denial of TVPA and IMF assistance to Tier Three countries may have the perverse effect of contributing to the socio-economic conditions that make trafficking lucrative in the first place.

The Millennium Challenge Account (MCA)

In 2002, President Bush announced a new program that attempts to diverge sharply from the traditional reliance on qualitative, in-house human rights assessment tools like the reports described above. This is the Millennium Challenge Account (MCA). The MCA allocates aid by a process of selectivity rather than conditionality, and it uses a set formula of quantitative indicators to guide decisions. The MCA's pool of grant money, which is separate from the funds already available through USAID, is to be disbursed to a small group of low-income countries that are selected on the basis of their

(USCIRF), which is also mandated to produce an annual report. The Commission report covers fewer countries, than the Office's report, and it is used to make policy recommendations to the executive and legislative branches of government.

³⁹ See US Department of State, *Victims of Violence and Protection Act of 2000: Trafficking in Persons Report*, 2004, <<http://www.state.gov/g/tip/rls/tiprpt/2004/>>.

⁴⁰ *Victims of Trafficking and Violence Protection Act of 2000*, 106th Cong., HR 3244: Section 103 (7).

⁴¹ *Ibid.*, Section 110 (a) and (c).

⁴² See for example Elizabeth Kelly, *Journeys of Jeopardy: A Review of Research on Trafficking in Women and Children in Europe* (Geneva: IOM Migration Research Series, No. 11, November 2002).

performance in three categories: “governing justly,” “investing in people,” and “promoting economic freedom.”⁴³

MCA Performance Criteria

Governing Justly:

- Civil liberties (Freedom House)
- Political rights (Freedom House)
- Voice and accountability (Kauffman-Kraay “Voice and Accountability” score)
- Government effectiveness (World Bank)
- Rule of law (World Bank)
- Control of corruption (World Bank)

Investing in People:

- Public primary education spending as percent of GDP (World Bank, national sources)
- Primary education completion rate (World Bank, national sources)
- Public expenditures on health as percent of GDP (World Bank, national sources)
- Immunization rates: DPT and measles (World Bank, UN, national sources)

Promoting Economic Freedom:

- Country credit rating (*Institutional Investor Magazine*)
- Inflation (IMF)
- 3-year budget deficit (IMF, national sources)
- Trade policy (Heritage Foundation)
- Regulatory quality (World Bank)
- Days to start a business (World Bank)

Based on information from the data-sources listed in parentheses above, countries are rated on their performance in these 16 areas. Countries that meet a certain threshold of performance in all three areas are eligible to apply for aid.⁴⁴ The chosen recipients—16 countries as of May 2005⁴⁵—are invited to set their own priorities for the aid, design project proposals, and set benchmarks for achievement. In addition to encouraging local

⁴³ See [MCA Data on the Web](http://www.mca.gov/countries/selection/data.shtml). 2005. Millennium Challenge Corporation.

<http://www.mca.gov/countries/selection/data.shtml>.

⁴⁴ The MCA plans to revise its selection criteria for FY 2006. For details of the proposed changes, see Steve Radelet, Bilal Siddiqi, and Mvemba Dizolele, “New Global Governance Indicators and the Possible Impact on MCA Qualification,” Center for Global Development, May 2005, <http://www.cgdev.org/doc/commentary/MCA%20projections%20FY%2006%20-%20FINAL.pdf>.

⁴⁵ The Center for Global Development keeps a running tally of eligible countries. For updates see <http://www.cgdev.org/section/initiatives/active/mcamonitor>.

ownership of development projects,⁴⁶ this collaborative approach is meant to ensure that the United States can hold receiving governments accountable for how aid is spent, and can increase or decrease future grants accordingly.

Most human rights activists would agree that how a government treats its citizenry is a relevant concern in discretionary aid allocations. After all, the majority of official development assistance is channeled through the receiving government's agencies, and it will never reach its intended beneficiaries if the government is more interested in lining its own pockets than in responding to the needs of its population. By this logic, the MCA's inclusion of human rights indicators in its governance-measurement schema is hardly controversial. Indeed, if actual disbursement rates live up to the MCA's promises, then it will represent a laudable increase in US bilateral grant aid. What is more controversial is which indicators, and which data sources, the MCA uses.

First, let us consider the MCA's choice of indicators. Insofar as it claims to depend upon "objective and quantifiable"⁴⁷ information from independent sources, the MCA is an attempt to answer the criticism that US aid decisions are inconsistent and politically motivated. To some extent, its standards-based methodology—relying on the same 16 quantitative indicators for each country—does succeed in insulating aid decisions from extrinsic diplomatic considerations.⁴⁸ Some critics argue that whether or not the selection process is consistent, the criteria themselves reveal such underlying biases (in favor of neo-liberal economic policies, for example) that any claims to objectivity are suspect. But this criticism misses the point of the entire endeavor: to promote the sorts of governments that conduct themselves in a way the United States considers beneficial for domestic and international stability.

A more relevant criticism could be that certain rights issues (like domestic violence or racial discrimination) are under-represented in the MCA's list of 16 criteria, as evidenced by the indicators' lack of disaggregation by gender, ethnicity, religion, or geography. This line of criticism is correct in pointing out that the MCA's indicators do not convey a sense of the distribution of rights and resources within a state, they lack the necessary balance of process and outcome that characterize rights-based indicators, and they fail to capture the impact of any actor other than the government. But, like the Kauffman-Kraay dataset, the MCA never claims to be a human rights measurement system *per se*; it is interested in rights not for their own sake, but instrumentally, for what they tell us about governance. Whether or not the United States needs to incorporate more rights-based indicators into its formula is something the human rights community

⁴⁶ The lack of transparency and local participation in traditional conditionality frameworks is coming under increasing criticism lately. The MCA is attempting to answer such criticism by involving receiving countries in policy-making processes. The World Bank's IDA eligibility assessments have been criticized along similar lines but has not yet made a move to enlist more participation from receiving governments; this was one of the problems raised by a 2004 review of the CPIA framework. See Shengman Zhang (Managing Director), "Country Policy and Institutional Assessments: An External Panel Review Panel Recommendations and Management Follow-up," report by a panel of external experts for the World Bank, June 2004. <<http://www.bothends.org/politiekcafe/polcaf18-country%20policy.pdf>>.

⁴⁷ *Millennium Challenge Act of 2003*, 108th Cong., 1st Session, SR 571: section 607 (a).

⁴⁸ Steven Radelet, "Bush and Foreign Aid" in *Foreign Affairs* 82:5 (September/October 2003) 104.

should debate; we will address this question in section V below. But their exclusion here may be a troubling conceptual gap, but it cannot truly be called an inconsistency in the MCA's methodology.

More obviously troubling methodological questions, though, should be raised regarding the MCA's choice of data sources. Because the human rights indicators in the "governing justly" category are drawn solely from the Kauffman-Kraay dataset and from Freedom House's *Freedom in the World* country rankings, they inherit the methodological problems of both. The weaknesses of the Kauffman-Kraay dataset have been described above: it focuses selectively on some rights issues and not others without a clear theoretical justification for doing so; it incorporates "human rights" data sources into a "governance" assessment framework without addressing the divergent aims of these two types of analysis; and it inappropriately uses non-time series data to support a time-series analysis. Freedom House's country rankings may also be questioned along theoretical and methodological lines, although, to their credit, they do make sure that any changes to their methodology are introduced gradually so that results from year to year will remain comparable. At any rate, the MCA's inclusion of both Kauffman-Kraay and Freedom House scores is redundant, since Freedom House's rankings have already been factored into the Kauffman-Kraay governance scores.⁴⁹

There are two main reasons why, despite their shortcomings, the MCA relies on these two particular data sources. One is that they are "external"—that is, they are not generated by the US government itself. This claim is slightly misleading, however: the Kauffman-Kraay dataset uses US State Department country reports as one of its human rights data sources, and Freedom House is largely funded by USAID. Another reason is that these sources offer broad geographical coverage over long periods of time. But this alone does not mean they are sufficiently sound sources on which to base aid determinations. If the MCA is an experiment worth continuing, its human rights components should be based upon better data, as well as a clearer conception of how human rights fit into a governance framework, and which human rights issues are truly relevant. The national and international human rights community would do well to involve itself in this effort—in refining indicators that capture what they mean to capture, and in helping to think more clearly about the link between human rights and governance.

V. The response of the human rights community: joining the debate

The World Bank and the United States may be the frontrunners in using quantitative measurement systems to guide their aid allocations, but they may not be alone for long. The Scandinavian countries are particularly keen to develop human rights and governance indicators, and the EU also appears to be contemplating a more

⁴⁹ For some thoughtful, technical critiques of the MCA methodology, see Steve Radelet, Sarah Lucas, and Rikhil Bhavani "A Comment on the Millennium Challenge Account Selection Process," Center for Global Development, 9 March 2004. <<http://cgdev.axion-it.net/docs/CGD%20Comment%20on%20MCA%20Selection%20Process.pdf>>.

standards-based approach to aid disbursements.⁵⁰ This presents a valuable opportunity for the human rights community to overcome its reticence about measurement and lend its voice to the debate.

Understandably, human rights actors have traditionally been reluctant to enter into discussions about human rights sanctions and aid conditionality. On the one hand, no human rights defender wants to reward an abusive regime by increasing its access to overseas funding and thus underwriting its hold on power. On the other hand, even systems of incentive-based selectivity like the MCA—just like systems of sanctions-based conditionality—can have the perverse effect of leaving the countries with the worst human rights records economically disengaged, and therefore less likely to pay heed to international pressure and improve their domestic practices. It is true that MCA grants are currently awarded *in addition* to (not instead of) other forms of assistance like USAID funding. Therefore, the US administration points out, ineligibility for MCA grants does not mean that a country is ineligible for *any* kind of official US assistance. But if MCA-style grants eventually began to displace other forms of ODA, then selectivity and conditionality would be effectively the same thing: countries that did not meet the MCA's standards in human rights and governance would have a smaller pool of alternative funding on which to rely. That is, “selectivity” can only truly be “selectivity” when there are other options available.

Even now, much more development assistance (of all sorts) is promised than is ever disbursed.⁵¹ Of the money that is disbursed, much of it is “tied”—meaning that goods and services related to dispatching that money must be purchased from US suppliers. This practice of imposing restrictions on ODA's uses that can drastically reduce the money's effectiveness.⁵² Thus, there is not as much money available outside

⁵⁰ Human rights conditionalities may become more formalized elements of EU aid policy in coming years. EuropeAid states that “the Commission has begun to consider progress in implementing institutional reforms for human rights, the rule of law and good governance as an element in defining allocations for individual countries.” The European Commission's statistical office, EUROSTAT, launched a project on measuring democracy, human rights, and governance in 2002. See Todd Landman and Julie Häusermann, “Map Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance,” report for the Statistical Office of the Commission of the European Communities (EUROSTAT). University of Essex: Human Rights Centre, 2003. <http://www.oecd.org/dataoecd/0/28/20755719.pdf>. It is still unclear how the indicators it has begun developing might come to guide EU aid allocation in the future. But, according to a document of 2001, the Commission intends “to work for a transparent approach to human rights and democratisation which is coherent and consistent between countries and regions and avoids double standards.” See “The European Union's Role in Promoting Human Rights and Democratisation in Third Countries,” communication from the Commission to the Council and the European Parliament COM (2001) 252, 8 May 2001, Annex 2. http://europa.eu.int/comm/external_relations/human_rights/doc/com01_252_en.pdf.

⁵¹ See “America, the Indifferent,” editorial, *New York Times*, 23 December 2004.

⁵² The OECD estimates that in 2002, the tying of aid reduced the value of global bilateral aid flows by \$5 billion. See UN, *World Economic and Social Survey: Financing for Development*, <http://www.un.org/esa/policy/wess/> at 121. For information of the tying of USAID funds, see Nancy Birdsall and Brian Deese, “Hard Currency: Unilateralism Doesn't Work for Foreign Aid, Either,” in *The Washington Monthly*, March 2004, 40. See also Steven Radelet, “Bush and Foreign Aid” in *Foreign Affairs* 82:5 (September/October 2003) 116.

the MCA as there would seem to be, if one were to judge solely by the amounts of money that had been pledged and allocated through USAID and other programs.

Finally, selectivity and conditionality alike can trigger herding responses among other donors, widening the gap between aid favorites and aid pariahs. The total amount of ODA available worldwide is of course not fixed, but neither is it infinite. If a small number of countries attracts the vast majority of the aid available from the major bilateral donors, regardless of whether these favored recipients have been identified positively (through selectivity) or negatively (through conditionality), the effect is the same: other countries will be neglected as donors, taking their cues from one another, consolidate their ODA on a few “best bets.” According to the UN, “since the 1980s, the top 20 [aid receiving] countries have received more than half of the net bilateral aid and [. . .] fewer than 50% of aid recipients have received 90% of all aid from [. . .] donors.”⁵³ Given this tendency for donors to mimic one another, the MCA’s “selectivity” approach may still leave the countries with the worst human rights records economically isolated, and therefore less likely to improve their domestic human rights practices.

The human rights community’s response to the ethical dilemmas of aid conditionality and selectivity has most often been to keep silent. Indeed, Amnesty International has never publicly addressed the fact that the MCA uses its country reports to determine countries’ eligibility for aid. This silence, while understandable, threatens to undermine the human rights movement—especially now, in 2005, as the UN calls for human rights to be considered a central component of the global poverty relief and development agenda.⁵⁴ It is high time for the human rights community to discuss what the relationship is between human rights and official development assistance. Embarking on such a discussion is necessary because human rights professionals are, at base, concerned with enhancing the wellbeing of people in the world, and ODA is a critical tool in the service of this goal.

At the very least, embarking on this discussion is necessary to ensure that our advocacy not be rendered ineffective. If, for example, the information contained in Amnesty International’s *Annual Reports* can be used to rationalize the withholding of ODA from countries that flout basic human rights standards, then the form of human rights promotion of which Amnesty has made such admirable use—international naming and shaming—may lose its effectiveness. The very governments who most depend upon external economic support are often, also, the governments with the poorest human rights records; when donor governments deny them aid on these grounds, they (and international CSOs as well) risk losing a key element of their diplomatic leverage on those borrowing governments.

⁵³ UN, *World Economic and Social Survey: Financing for Development*, <<http://www.un.org/esa/policy/wess/>> at 116.

⁵⁴ Kofi Annan has stressed that “[d]evelopment, human rights, and security go hand in hand. [. . . They are] all imperative; they also reinforce one another.” United Nations, *In Larger Freedom: Towards Development, Security, and Human Rights for All*, report of the Secretary-General, A/59/2005, 21 March 2005, paras. 14, 16, <<http://daccessdds.un.org/doc/UNDOC/GEN/N05/270/78/PDF/N0527078.pdf?OpenElement>>.

The trade-offs involved in such decisions demand careful attention. Perhaps decision-makers would be better served by qualitative assessments of the costs and benefits involved in each case than by standardized frameworks or statistical formulae. Proponents of the Millennium Challenge Account would argue that there *is* room for case-by-case consideration built into their system: countries whose scores land near the threshold of eligibility are judged individually, with an eye to relevant factors and extenuating circumstances. The question then becomes this: if the MCA's architects admit that the statistical index is insufficiently sensitive to determine the eligibility of these threshold cases, then why use it at all? Is it truly necessary as a time saving device? Some human rights advocates would argue that decisions about ODA are ineliminably political (not simply economic) and any system of eligibility-assessment should forthrightly reflect this fact, relying upon qualitative, country-specific information culled from local and external expert sources. This is, after all, what most of the World Bank and MCA data sources are before they are numerically coded and used to create ranking indices.

It may be that standardized eligibility frameworks are necessary to expedite the ODA determination process, at least in the preliminary stages. If this is the case, it may be that many members of the human rights community would prefer to take human rights norms out of these frameworks, acknowledging the complexity of such concepts and the inevitable sloppiness of ranking rights performance country by country. But it may also be that human rights activists think it is important to include human rights for consideration in ODA policy, and if so, the movement has some work to do to make sure the right things are included. It is now time to address these issues. The debate will go on, with or without the human rights community's participation.

Clarifying the concepts

Voices from the world of human rights can help to address several issues that have so far been neglected by donors. First, human rights practitioners and academics can help donors do some much-needed theorizing on what they consider "governance" to mean, and how human rights fit into this definition. They can encourage critical debate about whether, how, and why human rights measurement should be involved in aid disbursement. They can prompt donors to consider whether measures of governance that focus more on structure than on process, or more on stability than on justice, are missing something essential, and they can challenge one another on how best to bring their own human rights expertise to bear on these issues.

Constructing better indicators

Having worked through these conceptual questions, the human rights community should ensure that the indicators used in governance measurement tools reflect the conclusions reached in these debates. Where they find gaps in the current indicator sets, human rights actors could identify alternative, rights-based indicators that would help to illuminate the neglected aspects of countries' human rights situations, and argue for their inclusion.

Supplying better data

Data collection, too, needs to be critically reviewed according to human rights standards. Human rights organizations like Amnesty, Freedom House, and Reporters without Borders should seriously consider whether or not they are comfortable supplying data to governance measures. If so, they will need to address the methodological shortcomings outlined in this paper. If not, they should help to identify more appropriate data sources. The information collected by the UN through the treaty bodies, Common Country Assessments, national statistical bureaux, Special Rapporteurs and other bodies could be the best available source of reliable and systematic data.⁵⁵

Making the case to donor agencies

The human rights community's participation in these discussions may not be welcomed, though it is critically important. Naturally, the World Bank and the Millennium Challenge Corporation will not necessarily *want* to overhaul their indicators to make them more thoroughly rights-based. Donor governments and IFIs may argue that they have already incorporated rights to the extent that they need to—that for their purposes, it is enough to capture only those aspects of human rights that deal with stability, institutional capacity, and national aggregates. Whether this is truly enough information for donors to make sufficiently accurate risk assessments to guide their ODA “investments” is debatable—and the human rights community should join that debate. But certainly, the governance measures currently in use provide only a partial picture of the situation of which donors should be aware in aid-receiving countries. A broader snapshot such as rights-based indicators would be useful not only in determining countries' eligibility for aid, but also in tracking the effects of that aid once given. Because they capture both process and outcome, rights-based indicators—unlike most governance indicators—are designed to pick up on the effects of actors and duty-bearers other than the state. Donors may protest that, since aid is mainly channeled through government agencies in receiving countries, the only duty-bearer they need to evaluate is the government. They may not be interested in evaluating the behavior of militia groups, the labor practices of multinational corporations, or the prevalence of domestic violence behind closed doors so long as the government itself is not corrupt.

To this argument, the human rights community must reply: the receiving government is *not* the only relevant actor. These other duty-bearers are worth considering as well. But more importantly, you, donors, are also actors and duty-bearers, and you profoundly affect the countries to which you send aid. The impact of your grants or loans will be registered in the economic, social, and development sectors, of course; but your involvement can also trigger changes in the local rights situation. These changes will often be non-linear and difficult to predict without sensitive measurement strategies. If the World Bank gives Tajikistan a loan to help modernize its agricultural sector, for example, how will this affect Tajik children—on whose forced labor the cotton industry

⁵⁵ For more information on the information collected by the UN, see this author's other working paper, “The United Nations and the Measurement of Human Rights.”

depends? Will agricultural modernization lead to a decreasing reliance upon child labor and, thus, increases in school attendance and literacy rates, or will it instead lead to greater rates of child trafficking across borders or into other unregulated sectors as destitute parents lose a major source of income?⁵⁶

The current systems of governance measurement can hardly be expected to pick up on subtle changes in local rights situations like this, but it is extremely important that they do. If donors do not assess the human rights situation on the ground before, during, and after their engagement, they will not be able to track a crucial dimension of their own impact. For the sake of efficient and effective future planning, as well as for the sake of their moral obligations to the populations they select as their beneficiaries, donor governments and IFIs must be able to measure these effects. Donors need a map of the human rights landscape in the countries to which they give aid so that they can knowledgeably and responsibly plan their involvement there.

First steps toward donor accountability

Some donor governments⁵⁷ are beginning to acknowledge the importance of incorporating truly rights-based indicators into their measurement systems for the reasons outlined above. Britain's DFID⁵⁸ is working to produce diagnostic human rights

⁵⁶ A famous example of the unintended consequences that a donor's policies can have occurred in Bangladesh in 1992. When US Congressman Tom Harkin introduced a bill to Congress that would have prohibited the importing of goods made in Bangladesh with child labor. The bill was not passed, but the mere threat of sanctions triggered a quick response from Bangladesh. Employers in the garment industry dismissed about 50,000 child laborers. The government, however, made no provisions for their assistance upon dismissal, and a large number of these children found themselves in other forms of exploitative labor or servitude. See UNICEF, *The State of the World's Children*, 1997, "An Agreement in Bangladesh," <<http://www.unicef.org/sowc97/>>, at 60.

⁵⁷ In 1999, the Danish Centre for Human Rights commissioned a report entitled *Human Rights Indicators 2000*, which evaluates the governments of 70 countries by assessing their formal commitment to human rights and their *de facto* performance. See Hans-Otto Sano and Lone Lindholt, "Human Rights Indicators: Country Data and Methodology 2000" (Danish Centre for Human Rights) <<http://www.humanrights.dk/departments/international/PA/Concept/Indicato/Ind2000/>>. DCHR (now the Danish Institute for Human Rights) is not a government agency, although it was created by the Danish Parliament. DCHR's country scores are not meant to be used by the government in determining countries' aid eligibility. Instead, the report describes its usefulness as a contribution to human rights research in general and to DCHR's own annual planning in particular. The Danish Institute for Human Rights is also in the process of developing the Human Rights Compliance Assessment (HRCA), a diagnostic test that corporations can use to ensure that businesses comply with human rights standards. Based on Environmental Impact Assessments, the HRCA comprises about 1,000 human rights indicators and 300 questions based on the Universal Declaration of Human Rights, the primary ILO conventions, and other human rights declarations and conventions. See the Danish Institute of Human Rights, *Human Rights Compliance Assessment*, <<https://hrca.humanrightsbusiness.org/>>. See also Danish Centre for Human Rights, *Annual Report 2001*, <http://www.humanrights.dk/upload/application/ca1eeaad/beretn_eng.pdf>, at 23.

⁵⁸ In September 2004, DFID issued a draft paper recommending that the UK streamline its aid policies away from traditional aid conditionalities and toward standards that correspond to the UN Millennium Development Goals of poverty reduction, human development, and human rights. The terms of aid should be linked closely to benchmarks of success on poverty reduction, which will be determined according to the international MDGs and the receiving country's own poverty strategy. Conditions should take into account human rights and security as well. If the agreed conditions of aid are not observed—if, for example, the

measurement tools for use in assessing the human rights impact of its own programs. Sida, the Swedish aid agency, has produced a *Guide for Country Analysis from a Democratic Governance and Human Rights Perspective*, a checklist of basic human rights indicators to consider when setting its partner country development strategies.⁵⁹ Similarly, NORAD in Norway issued its *Handbook in Human Rights Assessment* in 2001,⁶⁰ which it describes as a tool for assessing the positive and negative effects of proposed overseas development projects. It includes a checklist of qualitative human rights questions that should be asked periodically in order to take stock of a project's impact on the local human rights situation. And finally, in March 2005, the Paris Declaration on Aid Effectiveness attempted to convene a multi-national conversation on this topic, drafting guidelines by which participating countries should assess the efficacy of their aid policies.⁶¹

These measurement initiatives are all still in early stages, and much work remains on refining the qualitative and quantitative indicators to be used in tracking effectiveness. But they do acknowledge the importance of assessing human rights not only for the sake of identifying countries “deserving” of aid, but also for the sake of better tailoring donor governments’ development strategies there. In using diagnostic tools to assess their own impact on local human rights conditions, donor governments accept a share of the accountability for human rights abroad, and they enable themselves to better plan and tailor their aid agendas to bring about the improvements that they—and the human rights community—identify as desirable.

VI. Conclusion

The United States and the World Bank operate on the assumption that what matters for ODA eligibility is whether receiving countries’ governments are politically stable and relatively responsive to the concerns of the majority of their citizens—and

receiving government is responsible for significant human rights violations, or if it refuses to allow progress toward the agreed benchmarks to be monitored—then aid may be suspended. These recommendations were adopted as policy in March 2005. See the UK policy paper, “Partnerships for Poverty Reduction: Rethinking Conditionality,” March 2005, <http://www.dfid.gov.uk/pubs/files/conditionality.pdf>.

⁵⁹ Sida, *Country Strategy Development: Guide for Country Analysis from a Democratic Governance and Human Rights Perspective* (no date). See <http://www.humanrights.se/svenska/Guide%20for%20Country%20Analysis-%20Sida.pdf>.

⁶⁰ Bård Anders Andreassen and Hans Otto-Sano, “What’s the Goal? What’s the Purpose? Observations on Human Rights Impact Assessment” Norwegian Center for Human Rights Research Notes 02/2004 (Oslo: Norwegian Centre for Human Rights, 2004). <http://www.humanrights.uio.no/forskning/publ/rn/2004/0204.pdf>.

⁶¹ According to the UN, the Paris Declaration ‘set out 5 major principles of aid effectiveness: (a) ownership of development strategies by partner countries, (b) alignment of donor support with those strategies, (c) harmonization of donor actions, (d) managing for results, and (e) mutual accountability of donors and partners. The Declaration also contained some 50 commitments to improve aid quality which were to be monitored by 12 indicators. Participants agreed to preliminary quantitative targets for only 5 of them, and the Declaration is particularly weak on commitment to improve alignment (no target for reliable recipient country systems, for coordinated donor capacity or for untying aid) and agreed to set targets for the other indicators by the Summit meeting of the General Assembly in September 2005.’ See UN, *World Economic and Social Survey: Financing for Development*, <http://www.un.org/esa/policy/wess/> at 120.

thus, whether they are likely to make productive use of loans and grants, contributing both to local poverty reduction and to global economic growth. These donors have therefore constructed quantitative governance measurement schema that aim to capture these criteria.

When these donors incorporate human rights concepts into their governance measurement schema, they do so selectively, setting standards, choosing indicators, and gathering data largely without the input of either local or international human rights actors. The donors we have examined here rely upon qualitative, external expert assessment sources, processing this data to produce numerical scores that can be used to arrive at composite country rankings. The human rights sources that they use currently include the reports of the US State Department, Amnesty International, Freedom House, and Reporters without Borders. Yet, as we have seen, these sources were not intended for—and may be ill suited to—the task being demanded of them. Their use raises methodological questions about redundancy, bias, and comparability over time, as well as broader, theoretical questions about whether it is possible (and indeed desirable) to incorporate human rights performance into standardized ranking indices for the purposes of guiding ODA policy.

Beyond these methodological and theoretical concerns, there are important conceptual problems involved in the ways human rights are currently being measured for the purposes of determining aid eligibility. This paper has argued that current methods of measuring rights may fail to reflect the very concerns that make the human rights perspective unique, and that differentiate it from a governance or development inquiry.

In short, the problems are as follows. The indicators which the World Bank and the MCA derive from these human rights data sources tend to focus on average country-wide performance rather than on the distribution of resources and remedies within a country (across geographic regions, ethnic groups, religious groups, and so on). They tend to emphasize stability over equity, structure and outcome over process, and aggregate welfare over the individual's experience. And finally, they account for only one duty-bearer—the government—while overlooking the human rights effects that other actors like militias, multinational corporations, and even the donors themselves can have. In so doing, these indicators lose much of what makes a rights-based inquiry unique and meaningful: concern with disparities among groups, concern with the complex interactions of rights-holders and duty-bearers, and concern with the processes involved in these interactions (as well as the legal and administrative structures in place and the average outcomes that are ultimately attained).

And so, indicator sets that claim to provide some measure of “human rights” lose much of what we in the field believe that term means. This is not simply a question of definitional coherence. Neglecting to evaluate all the interdependent aspects of the rights situation in a receiving country, donors may unwittingly exacerbate local problems. For example, ODA policies that are focused overwhelmingly on improving regime stability might succeed at the expense of ethnic and religious minority groups, and might stifle their claims to self-determination or cultural integrity. Indeed, without appropriate and

rights-sensitive measurement tools, donors may be unable to pick up on the unintended consequences that their own involvement and development assistance brings about.

For these reasons, a human rights perspective on governance measurement and aid conditionality is urgently needed. Given that donor agencies are *already* using human rights criteria in directing their official development assistance, the human rights community should lose no time in adding its voice, and its expertise, to the discussion. We must ensure that any human rights indicators used in determining aid allocations be conceptually clear, methodologically appropriate, and sensitive to the whole range of rights considerations that might be affected, for better or for worse, by official development assistance.