

WP 2005-02
February 2005



Working Paper

Department of Applied Economics and Management
Cornell University, Ithaca, New York 14853-7801 USA

REFORMING THE FORMULA: A Modest Proposal for Introducing Development Outcomes in IDA Allocation Procedures

Ravi Kanbur

It is the Policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

**Reforming the Formula:
A Modest Proposal for Introducing Development Outcomes in IDA Allocation
Procedures**

By

Ravi Kanbur*

www.people.cornell.edu/pages/sk145

**First Draft: October 2004
Second Draft: January 2005**

Contents

1. Introduction
2. The IDA Process and Formula
3. The Logic of the Formula, and a Critique
4. Outcomes Based Aid Allocation: Criticisms and Responses
5. Conclusion: A Modest Proposal

Abstract

This paper develops a modest proposal for introducing final outcome indicators in the IDA aid allocation formula. It starts with a review of the current formula and the rationale for it. It is argued that this formula, and in particular the Country Policy and Institutional Assessment (CPIA) part of it, implicitly relies too heavily on a uniform model of what works in development policy. Even if this model were valid "on average", the variations around the average make it an unreliable sole guide to the country-specific productivity of aid in achieving the final objectives of development. Rather, it is argued that changes in the actual outcomes on these final objectives could also be used as part of the allocation formula. A number of conceptual and operational objections to this position are considered and debated. The paper concludes that there is much to be gained by taking small steps in the direction of introducing outcome variables in the IDA formula, and assessing the experience of doing so in a few years' time.

* T.H. Lee Professor of World Affairs, International Professor of Applied Economics and Management, and Professor of Economics, Cornell University. Paper for presentation at the AFD-EUDN conference, Paris, November 25-27, 2004. The ideas in this paper have been presented at seminars and panels at Princeton University, IFAD (Rome), IFPRI (Washington, DC), the World Bank (Panel on Lessons of the 1990s), the DPRU/TIPS/Cornell conference on African Development (Cape Town), and at DFID's Conference on Reaching the Very Poorest (London). Parts of this paper draw on my contribution to the DFID conference, "What Change Does Attention to the Poorest Imply?" I am grateful to participants at these meetings for their helpful comments.

1. Introduction

How should aid donors allocate aid between recipient countries if their objective is to advance development?¹ This question poses both conceptual and operational issues. All donors have rules and procedures that feed into the determination of the level and composition of aid transfers to different recipients. In many cases there is an explicit formula which, while not determining in a mechanical sense, certainly sets the benchmarks from which the allocation decision begins. One such formula is the IDA allocation formula, but other donors have procedures that are similar in spirit.

A very simple framework would suggest the importance of two key factors in the allocation choice between potential recipient countries. First, how successful would this aid be in aiding development? Second, how is the development in one country to be valued against that in another? The first is an “aid productivity” question. The second is a “valuation of outcomes” question. The second question is relatively easy to answer if the donor’s valuation of development in recipient countries is clear. Given the development outcomes the donor is interested in, for example a reduction in infant mortality rates, a natural specification of the valuation is that a unit improvement should be valued more the worse is the starting point. Thus, roughly speaking, for any given degree of aid productivity, aid allocation should vary inversely with the level of development of a country (the exact relationship would need a closer specification of the valuation function).

¹ For overviews of the aid literature, see Tarp (2000) or Kanbur (2003).

The question on valuation of development outcomes is not without its complexities.² But it can be argued that, at least to some extent and especially in the wake of the consensus on the Millennium Development goals (MDGs), the international community has something of an idea of what it values as the outcome of development. Rather, it is the first question that has vexed aid analysts and practitioners alike, because the productivity of aid is not independent of the modalities of aid delivery and the usage of that aid. The arc of thinking has traversed a project oriented phase, where the outcomes of specific projects were the guide to aid allocation, and a policy oriented phase, where the policy parameters of the recipient country were seen as a better guide to the productivity of aid. The discussion has often been cast in terms of the much used, and abused, term “conditionality.”

At its most general, conditionality is nothing more than the rules and procedures according to which a donor transfers resources to a recipient. To be against conditionality in general doesn't make sense. The devil really is in the detail—the detail of the rules and procedures according to which aid is allocated and disbursed.³ And these rules and procedures kick in at different levels, in the overall resource envelope allocated to a country, in the division of this envelope between different types of assistance, for example project or program modalities, and in the specific conditions that apply to particular projects or programs.

² See Kanbur (2004a)

³ For a discussion of conditionality in the context of the history of development assistance, see Kanbur (2003).

This paper is about the logic used in deciding the allocation of the overall aid resource envelope for a country. Since total resources are finite, such allocation has to be based, explicitly or implicitly, on a comparison of relevant features of different recipient countries. Perhaps the most prominent such method for comparison is the IDA allocation formula, not simply in terms of the total volume of resources that are allocated but because it is generally recognized that IDA procedures have a strong influence on the procedures of other donors as well. The component that is of specific interest in this paper is the method of cross-country comparison, the Country Policy and Institutional Assessment (CPIA) formula. The paper considers the logic of this formula, and proposes a revision to it.⁴

The plan of the paper is as follows. Section 2 outlines the IDA allocation procedure and the role of the CPIA in this procedure. Section 3 discusses the logic behind the use of the CPIA and offers a critique. Section 4 proposes allocations based on development outcomes and debates the major criticisms of this approach. Section 5 concludes by offering a modest revision of the CPIA as the first step to moving towards a development outcomes based approach.

⁴ There are, of course, many aspects of the development assistance process that are important but are not covered in this paper, for example, the sometimes perverse incentives in aid agencies to move money rather than focus on the best use of that money, or the interplay between foreign policy objectives and development objectives in the realpolitik of development assistance allocations. Also, my specific focus is on IDA, so I will not be discussing formulae used by other agencies such as the European Union, DFID or USAID.

2. Outline of the IDA Formula⁵

At the core of the logic of the IDA allocation process is a balance between “needs” and “performance”. Needs are measured straightforwardly by national income per capita, GNIPC. Performance is measured by a performance rating, PR, which is the focus of this paper. The allocation per capita for a country is a function of GNIPC and PR. In fact, the specific relationship is (World Bank 2003a):

$$\text{Allocation per capita} = f(PR^{2.0}, GNIPC^{-0.125})$$

Thus the performance rating is raised to the square power and per capita income is raised to a negative power, minus 0.125, and these two are then combined to decide the allocation. The function $f()$ is chosen to reflect the fact individual country allocations have to add up to the total resources available. A feature to note is that the performance rating has a much higher weight than the measure of needs. But this is not our major concern in this paper. Rather, the focus is on how the PR index is constructed and the logic behind this construction.

Before turning to the PR index, some further clarifications on how the above formula is used. The allocation per capita derived above is not a hard and fixed amount, but rather a “norm”. The detailed determination of the allocation, and of the composition of this

⁵ The procedures and the formula are summarized in World Bank (2003a) World Bank (2003b) and updated in World Bank (2004a).

allocation between different types of assistance, is done in the Country Assistance Strategy (CAS). To quote World Bank (2003a):

“The allocation norm establishes the financial resources available for each IDA country for the following three fiscal years. The allocation sets the resource envelope that each country could expect to receive if its performance stays the same and assuming a pipeline of quality projects -- but is not an entitlement. In the case of a new CAS the allocation norm will set the base-case financing scenario....The CAS financing scenarios may be adjusted to reflect special country circumstances, which will be spelled out in the CAS.” (World Bank, 2003a, p2).

Moreover, there are a number of exceptions to the norm derived above:

“In addition to their performance-based allocations, all countries are allotted a basic allocation of SDR 3 million (about US \$ 4 million). In terms of per capita allocations, this benefits in particular the small states. There are some important considerations that merit exceptions to the allocation norms. First, “blend” countries with access, or potential access, to IBRD receive less than their norm allocation due to their broader financing options. Second, post-conflict countries can, when appropriate, be provided with additional resources in support of their recovery and in recognition of a period of exceptional need. And third, additional allocations may be provided in the aftermath of major natural disasters.” (World Bank, 2003a, p2).

However, despite these caveats, the allocation norm, and the performance rating that underlies it, is a central feature of the whole process.

How is the PR index derived? At the heart of it is the Country Policy and Institutional Assessment (CPIA). The procedure for 2003 is as follows (the 2004 procedure has some changes that are noted below). Essentially, this is an assessment of a

country on each of twenty items divided into four categories, as shown in Table 1. Each of these items is then scored by Bank staff on a scale from 1 (low) through 6 (high). The broad interpretations of these scores are given in Table 2. The specific guidelines are elaborated in the 2003 CPIA questionnaire:

“Countries should be rated on their current status in relation to these guidelines and to the benchmark countries in each region, for which the agreed ratings have been provided to the staff. Please assess the countries on the basis of their currently observable policies, and *not* on the amount of improvement since last year *nor* on intentions for future change, unless the latter are virtually in place.... As described in these guidelines, a “5” rating corresponds to a status that is good today. If this level has been sustained for three or more years, a “6” is warranted, signifying a proven commitment to and support for the policy. Similarly, a “2” rating represents a thoroughly unsatisfactory situation today. A “1” rating signifies that this low level has persisted for three or more years, and therefore that the resulting problems are likely to be more entrenched and intractable.” (World Bank, 2003b, pp 1-2.)

Finally, a simple unweighted average of these scores is taken to give the CPIA index. Individual country scores are not released to the public, only country quintiles are made available (this is slated to change in 2005). The results for 2003 are given in Table 3.

Before turning to the specific categories and the scoring criteria for them, it is worth specifying how exactly the CPIA feeds into the PR. First the CPIA is combined with the Bank’s Annual Review of Portfolio Performance (ARPP), the weights being 80% for CIPA, 20% for ARPP. Then this weighted average is multiplied by a “governance factor”. The governance factor is built up as follows. First, an unweighted average is taken of the scores for six governance-related criteria in the CPIA, #4 and #16-20 (see Table 1), and of

a seventh score, on the “procurement practices” criterion from the ARPP assessment process (since it is not the focus in this paper, the ARPP process is not discussed in any further detail). This average score is then divided by 3.5 (the mid-point of the 1-6 scoring range), and this ratio is raised to the power of 1.5. This procedure effectively ends up giving significantly greater weight overall to the governance criteria in the CPIA. (Note that this is the procedure for 2003. For 2004, a revised procedure was adopted, as set out in World Bank, 2004a).

The components of the CPIA are thus central building blocks in the whole process. There are specific guidelines for the scoring of each of the 20 items that make up the CPIA. Tables 4, 5, 6 and 7 lay out these guidelines for one component from each of the four major categories in the CPIA: Fiscal Policy under Economic Management, Trade Policy and Foreign Exchange Regime under Structural Policies, Equity of Public Resource Use under Policies for Social Inclusion/Equity, and Transparency, Accountability and Corruption in the Public Sector under Public Sector Management and Institutions. Note that guidelines are specified only for scores of 2 (unsatisfactory), 3 (moderately unsatisfactory), 4 (moderately satisfactory), 5 (good); a score of 1 is simply “unsatisfactory for an extended period” and a score of 6 is “good for an extended period”.

Finally, we note that in 2004 certain changes to the CPIA process were accepted by World Bank management (see World Bank, 2004a). Among these are to disclose CPIA scores from 2005 onwards and to establish an independent expert standing committee to review the CPIA methodology every three years. These movements are greatly to be

welcomed. In addition, the governance factor calculation was changed, and the number of CPIA categories was reduced to 16, as given in Table 8. However, albeit with new categories, and a new procedure for calculating the governance factor, the essence of the CPIA method and the IDA allocation formula are left unchanged.

This completes the outline description of the IDA formula, and its centerpiece, the CPIA scores. What is the logic underlying this method of aid allocation? We turn now to this question.

3. The Logic of the Formula, and a Critique

There are many specific and operational criticisms of the IDA allocation process. The CPIA is done behind closed doors by Bank staff, with little or no scrutiny from outside independent observers (slated to change in 2005). The ARPP remains an under scrutinized assessment procedure, linked as it is to internal Bank procedures. The way the “governance factor” enters the formula is convoluted at best. And it is not all clear where the different weights and exponents used in various parts of the formula come from. Why, for example, is PR raised to the power 2, while the governance score ratio is raised to a power of 1.5 to give the governance factor? Why exactly is GNIPC raised to the power of minus 0.125? But the main concern in this paper is not with these specifics—any formula will have to make such operational specifications and defend them the best it can. Rather, our concern is with the fundamental logic of the process.

As noted in the introduction, any logic for allocating development assistance resources to a poor country must have two components—how much the assistance can be translated into improvements in outcomes that the donor cares about (“aid productivity” or “performance”), and how much the donor values these improvements in outcomes (“need”). Thus if D is a measures of the final development outcomes and W(D) is the donor’s valuation of it, then the impact of aid A can be written mathematically as:

$$dW/dA = [\delta W/\delta D] \times [\delta D/\delta A] = \text{Need For Aid} \times \text{Productivity of Aid} \quad (1)$$

The first term on the right hand side values development outcomes as seen by the donor, while the second term measures the impact of a unit of aid on development outcomes, in other words, its productivity. If the value the donor places on the outcomes declines as the outcome improves, then the need dimension can be captured by an inverse function of the level of the desired variable. In the IDA allocation formula this is done simply by taking the per capita national income of a country and raising it to the power of minus 0.125.

Wealthier countries will get lower allocations through this component of the formula. Thus the IDA formula essentially captures need through the income criterion, and does not go directly to indicators such as infant mortality, maternal mortality, girls' education and other components of the MDGs, through which the international community has presumably expressed its objectives of the development process—the outcome variables that it is interested in. However, we will set this aspect of the IDA formula to one side, since the main focus in this paper is on the way that performance is measured.

Conceptually, if we hold the needs part of the formula constant then more aid should flow where its impact on objectives is greatest. If we could identify environments where aid productivity is highest, in other words where improvements in final development outcomes of interest, per unit of aid flow, would be greatest, then more aid should be allocated to those environments. Presumably the performance rating part of the IDA formula, and specifically the CPIA component of it, attempts to identify high aid productivity environments. The logic must be that a higher score on any of the twenty components of the CPIA enhances the productivity of aid flow and therefore argues for greater aid flow. These scores are then aggregated with equal weights cross the twenty

categories. There are two possible logics behind this last step. The argument could be that the twenty categories are equally valuable to aid productivity, or the argument could be that we have no information on the relative contribution of each category to overall aid productivity so, on the principle of insufficient reason, each category should be given equal weight.

But perhaps the most striking yet least remarked upon feature of the PR formula, and especially the CPIA part of it, is that *it is the same for every country*. The twenty categories are the same for each country, the guidelines for what gets a high score in each category are the same for every country, and the weighting scheme across the twenty scores (equal weighting) is the same for every country. What could be the logic behind this uniformity in country treatment?

One way to uncover the logic is to consider the literature on “cross-country growth regressions”, not least because this literature has had a tremendous influence in thinking on development strategies and aid strategies. In this literature, economic growth in a country is seen as a function of a number of determining variables. If growth rate is G then growth in country i is given by:

$$G_i = \alpha + \beta Y_i + \theta X_i + \gamma A_i + \eta X_i A_i + \varepsilon_i \quad (2)$$

where Y is a vector of structural variables that the government cannot control (such as a country’s geography and climate), X is vector of policy variables (like fiscal deficit, tariffs,

percentage of government expenditure devoted to primary education, or independence of judiciary) and ε is a classical stochastic error term. The coefficients in α , β , θ , and η translate the impact of their respect variables to growth. Thus according to the world view implicit in equation (2), a country's growth depends upon structural features that the government cannot control, policy variables that the government can control, aid flows, and an interaction term between aid flows and these same policy variables. This is, sometimes quite literally, the family of regressions that have been run over and over again in the literature, including the well known contribution of Dollar and Burnside (2000), and subsequent large numbers of papers by other authors.⁶ While a relationship like (2) is most often estimated for growth as the variable to be explained, there is no reason why in regression analysis the dependent variable cannot be, as it sometimes is, another development outcome variable like the infant mortality rate or life expectancy. Then we would simply replace G_i with D_i .

There are a host of data and econometric problems associated with estimating an equation like (2), but they are not my main focus and I want to set those aside for now. But one point to emphasize is that (2) sees no role for the aid flows themselves to influence policy, in other words, it sees no role for conditionality in changing government policy. This is surely right, because if the experience of two decades has taught us anything, it is that the development assistance tail cannot wag the domestic political economy dog. Rather, we should take the policies as emerging out of the domestic political economy, and

⁶ See, for example, Hansen and Tarp (2000), Dalgaard and Hansen (2001), Guillaumont and Chauvet (2001), Easterly, Levine and Roodman (2003). See also the survey in Kanbur (2003).

take them as givens in the aid allocation decision.⁷ If a relationship like (2) does indeed hold in a cross section of countries, then the implications for aid allocation are clear, since in mathematical expectation,

$$\delta G_i / \delta A_i = \gamma + \eta X_i \quad (3)$$

The productivity of aid is then given by the right hand side of (3), and we see there the values of the policy variables in country i , the elements of the vector X_i , weighted by the elements of the vector η , which are estimated from the regression run on cross-country data. The logic of the IDA formula is now clear. Equation (2), and its derivative, equation (3), lead to the scoring function given in the right hand side of (3). This tells you which policies should be counted (the elements of X_i) and how they should be weighted (the elements of η).

Having laid bare the logic, let us consider it further. The basic point is that the scoring rule in (3) is only as valid as the underlying model in (2) and its econometric estimation. First of all, it is not clear that anybody has ever run a regression with the twenty policy categories in the IDA formula, using the scores for policies as exemplified in Tables 4-7 to generate the elements of X_i , and even if they did, it is almost certain that they would not get a result where the elements of η were all equal, thereby giving the equal weighting rule. Rather, the categories in the IDA formula reflect an accretion of factors

⁷ In fact, in Burnside and Dollar (2000) a jointly estimated equation testing for the impact of aid on policies finds no such relationship. On conditionality, there is of course a huge literature. For example, see Guillaumont and Guillaumont (1995), Kanbur (2000, 2003), and Adam et.al (2003).

thought to be important to the development process, under different arguments made in different contexts. The equal weighting of the different factors then really does reflect the “principle of insufficient reason”, rather than a reasoned logic leading to particular combination of key policy factors that impact on the productivity of aid.

But perhaps most important is the fact that a common scoring rule for all countries, which is what (3) is, depends upon a common development outcomes model, for all countries. Put another way, Equation (2) effectively assumes that *all* factors explaining outcome variations across countries have been successfully accounted for in the variable included in equation 1. And the effect of an explanatory variable on development outcomes is *identical* across countries. Any variation across countries over and above that accounted for by the explanatory variables is purely random, not amenable to further parsing.

Over the past decade, dissatisfaction has been growing with the estimation of a cross-country “average relationship” leading to “best practice” policy guidelines which are common to all countries. This view, that variations around the estimates of average relationships like (2) are not simply pure random variations, but reflect country specific factors that are not captured in our model and in our data, is powerfully put in a recent report from the World Bank itself, *Economic Growth in the 1990s: Learning from a Decade of Reform*:

“The Study concludes that valid general principles do not imply generic “best practice” policy or institutional solutions....

Regarding macroeconomic policies for example, the findings emphasize the importance of institutions underlying macroeconomic stabilization, the risks associated with external financial liberalization, the disruptions associated with episodes of exchange rate appreciation, and the sometimes excessive focus on minimizing inflation in the short term....

Regarding trade, the analysis highlights the fact that countries that have successfully integrated into the world economy have followed different approaches and also adopted a range of complementary policies, making it difficult to pin down the exact relationship between trade integration and growth....

Perhaps the lesson of the lessons of the 1990s is that we need to get away from formulae and realize that economic policies and institutional reforms need to address whatever is the binding constraint on growth, at the right time, in the right manner, in the right sequence, instead of addressing any constraint at any time....” (World Bank, 2004b, pp vi-vii).

No doubt this view, put forward by one team writing one report, will be debated heatedly within the World Bank and without. But let me record here that I support the “end of certainty” heralded by this report. In this context, then, what I want to highlight again is that the CPIA does not contain any final outcome variables like poverty, extreme poverty, girls’ enrollment, maternal mortality rates, infant mortality rates etc. What it has instead is a series of intermediate variables like trade policy, regulatory policy, property rights, corruption, etc, which we hope will eventually influence the outcomes we are truly interested in. In effect, it has an implicit model of the development process which says that if the scores on the categories in the CPIA improve, then development outcomes will improve, or rather, the productivity of aid will improve. Over the years these categories have broadened and increased to 20, and then most recently decreased to 16, but the basic logic that we have the right and complete model, captured in the CPIA, has not changed.

And this is then the model that is implicitly assumed to be valid for every country to which the aid allocation formula applies.

My contention is that the evidentiary basis for imposing across countries this implicit common model of the development process that supposedly leads to improvement in final outcomes, is weak.⁸ It is weak for growth, and it is weak for development outcomes. Despite the support of the World Bank (2004b) report and the analysis therein, this may be interpreted as a controversial view, so let me be clear about what I am saying. I am not necessarily questioning that the model implicit in the CPIA is a good representation of the average across countries, although in other contexts I would question this as well. Rather, what I am saying is that the country variations around this average, deriving from myriad country specificities that cannot be captured by outside data and outside observers, are large and complex. How else can we explain the fact that Bangladesh, a country that far outstrips its comparators on improvement in social indicators in the last decade, is nevertheless at the top or near the top of Transparency International's corruption index? How else can we explain that once the fiscal deficit is in a range of, say, 2 percent of GDP, further reductions do not necessarily contribute to increased investment and growth? How else can we explain the fact that two countries can spend about the same amount on primary education, yet one country has higher enrollment rates and test scores? These variations are not due to random factors, but specific local factors that are not captured and perhaps cannot be fully captured in our models. The problem lies not in estimating an average relationship given the data that we have; the problem lies in then using this average relationship to make country specific judgments.

⁸ Apart from World Bank (2004b), see the many references in Kanbur (2004a).

Lest I am misunderstood, let me clarify further. There are certainly extreme situations, like hyperinflation, a double digit fiscal deficit ratio, a trade system rife with mutually inconsistent quantity controls, a production sector dominated by highly inefficient state enterprises, extremely low spending on education and health, etc, where general prescriptions about the direction of movement are indeed valid, although even here there may be pace and sequencing issues. But in “normal” cases lessons drawn from the average relationships may well obscure the local specificities that determine the success of policies and interventions.

So, in the face of these critiques of the underlying logic of the CPIA’s role in the allocation of IDA across countries, what is one to do? The next section turns to this question.

4. Outcomes Based Aid Allocation: Criticisms and Responses

If the average cross-country relationships that underlie the CPIA procedure, and thus the IDA allocation formula, cannot truly capture the cross country variations in the productivity of aid, what is the alternative? One alternative is of course to try even harder to capture country specific variations in estimating the relationship (2), with better data and more elaborate methods. This is a fruitful line of enquiry, but it is not my focus. Such further detailed investigation may never incorporate all country specific variations. In any event, while that exercise is ongoing, we have a specific question about what to do with the IDA formula.

The answer I propose, in an extreme and pure form to begin with, follows directly from expression (1), and is as follows. The needs side of the story can be measured by the levels of the outcomes we are interested in, but the performance side can be measured by the rate of improvement of these very same outcome variables over a given period of time up to the point of assessment, suitably normalized by the total aid flow over this period. A country that has very low levels of girls' enrollment in primary schools should get more aid on grounds of need. But a country that is showing rapid improvements of girls' enrollment from this low level, relative to the aid it is receiving, should get even more. A country that is showing relatively slow rates of improvement should get relatively less on account of this measure of performance.

In a purely conceptual sense I am arguing for an aid allocation formula that depends solely on outcomes for the very poor—levels to measure needs, change to measure performance.⁹ There are of course many criticisms, conceptual and operational, of this position. I want to rehearse a number of these, and present some responses.

**Which outcome variables will be chosen?*

This is an important question, and focuses the donor's attention on the true underlying objectives. For all its problems, the MDG process has begun to forge an international consensus on these outcome variables. But the lack of complete consensus on outcome variables seems to be an odd reason to move to intermediate variables—which have relevance presumably only because they will affect the outcomes we want favorably. If we do not know the outcomes we want, how can we know the intermediate variables that give rise to these outcomes?

**How will the different outcome variables be combined together?*

Another good question. This needs further discussion and consensus. But note that the 20 (or 16 from 2004 onwards) CPIA scores are also aggregated together to form one index, with little justification for the method of aggregation (a simple unweighted sum).

⁹ The idea of using actual outcomes in guiding aid allocations is not at all new, of course. For an argument that leads to this proposition, from the earlier literature on the design of conditionality in adjustment lending, see Collier et. al. (1997).

**What about the “start up” problem? A country coming out of conflict, for example, will be penalized because it cannot show a track record of performance on outcome variables.*

This is a generic issue no matter how we judge performance. As noted in Section 2, the current IDA allocation process does indeed make special provisions for post-conflict countries. There is no reason why these principles cannot continue to be used in the outcome based approach. For example, other criteria could be used for a set period in these cases.

**What about shocks and random events that can give a bad performance (or a good performance) in a year, independently of the government’s actions?*

Such variations can be handled in principle by taking time averages of the outcome variables—the averaging can be statistically as sophisticated as warranted. It should be noted that shocks often are factored into the subjective judgments that go into the assessments of the sub-categories of the CPIA (for example, if the tariff reductions did not go as far as might have been hoped because of a revenue crisis stemming from a decline in commodity prices). Moreover, allowances are made for natural disasters and the like in the current CPIA and similar procedures can surely be maintained.

**With this “backward looking” assessment of performance, can an incoming government not “live off” the achievements of its predecessor? Should we not be holding each government accountable for its actions, and indeed trying to influence these?*

These questions embody the twin conceits that we from the outside know the specificities of what needs to be done, so we can then judge actions, and that the tail of outside development assistance can wag the dog of domestic political economy. It should be clear from the discussion in the previous section that I am skeptical of both of these propositions.

**What about the reverse problem, of a government that comes in after years not of conflict and collapse, but simple policy neglect of the poorest, and now wants to take this on as serious issue? There is no track record of performance on outcomes, but the government wants to take actions which it claims will lead to improved outcomes—is this not another “start up” problem that deserves external support?*

I am sympathetic to this argument (note that it holds just as well with the current CPIA methodology), but at the same time I am wary of repeated “new dawn” arguments. But surely we can devise mechanisms to handle these cases by ring fencing limited funds for start up and then letting performance measurement on outcomes take over, just as we seem to do for the “new dawn” cases under the current methodology.

**What about data quality and availability for outcome variables? Is this not notoriously bad? How can we possibly run an aid allocation system based on such incomplete data?*

These are all valid points. But, first, data for the current intermediate variables (e.g. property rights) are not problem free either. Second, most importantly, we should be investing in monitoring of variables that we are truly interested in. It is odd that we invest in collecting information in intermediate variables, but not on the outcome variables of ultimate interest. But of course one of the reasons we do so is because the intermediate variables enter the aid allocation formulae! This cycle has to be broken somewhere. Indeed, assistance for monitoring development outcomes for the poorest can be (almost) unconditional in any setting. Interestingly, criterion #15 of the 2003 CPIA is “Monitoring and Analysis of Poverty Outcomes and Impacts.” IDA is currently emphasizing measurement of development outcomes. Although this exercise is currently seen as being outside of the CPIA process and IDA aid allocation process (indeed, monitoring as a separate criterion is not one of the 16 categories in the new 2004 VPIA—see Table 8), the emphasis on measurement and monitoring can only help if the objective is to make aid allocations more outcome based.

**If we make aid allocation depend on performance measured by outcome variables, will there not be an incentive for the government to manipulate this information?*

But this is true of any measure on which aid allocation will be based. Measures of Domestic Credit Creation (a classic intermediate variable), for example, are not free of the possibility of manipulation. This merely strengthens the argument for strong, locally based mechanisms of monitoring and evaluation.

**Will not the focus just on the outcomes detract from an understanding of the development process—do we not need the intermediate variables for that?*

Nothing I have said stops us from analyzing the development process, developing models and learning the lessons from country experience. And, indeed, we can convey those lessons from one country, say, to other countries. But what the focus on outcomes does is that it prevents the easy temptation of saying that because X worked in country A, we will condition aid on X being done in country B. In fact, it may be that country B tries Y and gets the same or better result. If two countries are delivering the same improvement in outcomes using very different intermediate methods, that should be interesting to us as analysts, but as aid allocators there is a strong argument for treating the two countries the same.

Where, then, do the criticism and responses leave us? The next section proposes a resolution.

5. Conclusion: A Modest Proposal

Thus, many of the criticisms of the pure proposal, to make the performance part of IDA's aid allocation formula based solely on improvements in outcomes per unit of aid flow, apply equally to the current IDA process, and there are reasonable responses to all of the criticisms. But the criticisms are nevertheless powerful, and suggest a carefully modulated move in the direction of outcome based allocation, learning and improving as we go along.

My proposal is therefore as follows.

While leaving the current IDA allocation methodology essentially intact, IDA should introduce one new category of scoring in the CPIA. This category should evaluate the evolution of an actual development outcome variable up to the present. The choice of variable is open. It will depend on international consensus and on data availability considerations, but surely the elements of the MDGs are likely candidates. Once an outcome indicator is chosen, its evolution should be described using appropriate time series averaging techniques. The rate of change of the averaged time series, per unit of aid flow, should then be scored in the same way as the other CPIA categories, from 1 to 6. Guidelines should be developed for this scoring, just as guidelines were developed for the current CPIA categories. The equal weighting procedure would still be maintained to give the CPIA score. All other aspects of the current IDA allocation procedure would be left unchanged by this proposal.

I hope that this proposal is seen as one that is practical and achievable. It responds to basic concerns about the logic of the CPIA and the IDA allocation formula, and yet it is not a radical departure from the current procedure. But to achieve it will require a fair amount of technical work to lay the background for choice of indicator, for the averaging technique, and for the guidelines on scoring. This work can be done, but it will require that we start soon, in time for the next CPIA in 2005. After three years, the experience can be assessed, and the possibilities for introducing new outcome variables (or indeed dropping them altogether) can be evaluated. The fact that there is to be a standing expert committee on the CPIA (see World Bank 2004b) means that this work can be guided by that committee, which can also facilitate technical discussions and consensus building. I commend this modest proposal to the World Bank.

Table 1

2003 CPIA Categories

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 (2003b)

Table 2

2003 CPIA Ratings Scale

1 (low) through **6** (high)

1 Unsatisfactory for an extended period

2 Unsatisfactory

3 Moderately Unsatisfactory

4 Moderately Satisfactory

5 Good

6 Good for an extended period

Intermediate scores of 2.5, 3.5 and 4.5 may also be given.

Scores of 1.5 and 5.5 may not be given.

Source: World Bank 2003(b)

Table 3

2003 IDA Country Performance Ratings

First Quintile Armenia, Benin, Bhutan, Cape Verde, Ghana, Grenada, India, Madagascar, Mauritania, Samoa, Sri Lanka, St. Lucia, St. Vincent and the Grenadines, Tanzania, Uganda

Second Quintile Azerbaijan, Bosnia and Herzegovina, Burkina Faso, Dominica, Guyana, Honduras, Malawi, Mali, Mozambique, Nepal, Nicaragua, Pakistan, Rwanda, Senegal, Serbia and Montenegro, Republic of Yemen

Third Quintile Albania, Bangladesh, Bolivia, Cameroon, Ethiopia, The Gambia, Indonesia, Kenya, Kyrgyz Republic, Lesotho, Maldives, Moldova, Mongolia, Vietnam, Zambia

Fourth Quintile Burundi, Chad, Democratic Republic of Congo, Republic of Congo, Côte d'Ivoire, Djibouti, Eritrea, Georgia, Guinea, Kiribati, Lao PDR, Niger, Sao Tome and Principe, Sierra Leone, Tajikistan

Fifth Quintile Angola, Cambodia, Central African Republic, Comoros, Guinea-Bissau, Haiti, Nigeria, Papua New Guinea, Solomon Islands, Sudan, Togo, Tonga, Uzbekistan, Vanuatu, Zimbabwe

Countries not rated in 2003 exercise: Afghanistan, Liberia, Myanmar, Somalia, and Timor-Leste.

Note: Number of countries per quintile varies due to equal scores at cut-off levels.

Source: World Bank (2003c)

Table 4

2003 CPIA: Score Guidelines for Fiscal Policy

- 2** Fiscal balance will likely lead (or is already leading) to inflationary financing, crowding out of private sector investment, an unsustainable current account deficit or an unsustainable level of public debt; or fiscal policy is not making a serious attempt at provision of public services and infrastructure essential to growth.
- 3** Sporadic efforts at macroeconomic stabilization through fiscal policy but not maintained consistently or implemented through temporary measures like ludicrously low real public sector wages or cuts in projects or services with high long run returns; or attempts at public services and infrastructure provision are sporadic and concentrated in not very cost-effective uses of funds.
- 4** Consistent maintenance of macroeconomic stability and fiscal sustainability through appropriate levels of the fiscal balance and lasting adjustment measures with only occasional lapses. Public service provision is good in some sectors but still inadequate in others.
- 5** Fiscal policies are consistent with adequate provision of high quality public services and infrastructure for economic growth and generate a fiscal balance that can be financed (including with aid flows where applicable) in a non-inflationary way and is consistent with adequate credit for the private sector and a sustainable path of public debt.

Source: World Bank 2003(b)

Table 5

2003 CPIA: Score Guidelines for Trade Policy and Foreign Exchange Regime

2 Average tariff (weighted by global trade flows) is high (over 30%). High and erratic import and/or export barriers, including quantitative restrictions and/or state trading monopolies. Export taxes or quantitative restrictions frequently used. Customs or political authorities make discriminatory or *ad hoc* exemptions. Valuation procedures arbitrary and artificial exchange rates result in substantial over or under valuation of goods for customs purposes. Clearance of goods requires many approvals, arbitrary fines, frequent bribes to customs officials and involves long delays. Foreign exchange rationed or an administered foreign exchange regime with multiple exchange rates.

3 Average tariff 20-30 percent. Coverage of quantitative restrictions reduced to 15 per cent or lower. Export restrictions mostly phased out. Duty exemptions frequently used to offset the adverse impacts of import barriers on inputs used in production of exports or for approved investment projects. Foreign exchange convertible for most current account purposes. Customs clearance involves high inspection ratios for imports, and interaction between officials and importers encourages corruption.

4 Average tariff 10-20 percent. Quantitative restrictions cover only a very small percentage of imports (under 5 percent). Virtually no export restrictions. Duty exemptions provided only in accordance with well-defined rules. Foreign exchange convertible for virtually all current account purposes. Customs procedures are well-defined, quick, efficient and impersonal, and staff professional, although some “tea money” payments to expedite clearance may still be present.

5 Average tariff (weighted by global trade flows) is low (10% or less), with low dispersion and insignificant or no quantitative restrictions or export taxes. Trading monopolies absent or unimportant. Indirect taxes (e.g. sales, excise, surcharges) do not discriminate against imports or exports. Efficient and rule-bound customs administration. IMF Article 8 status. Minimal or no foreign exchange restrictions on long-term investment capital inflows.

Source: World Bank 2003(b)

Table 6

2003 CPIA Score Guidelines for Equity of Public Resource Use

2 Most public expenditures for economic and social services do not benefit the poor more than the better off. The government has not identified individuals, groups or localities that are poor, vulnerable, or have unequal access to services and opportunities, does not have appropriate programs, and has no plans. Spending on economic and social services targeted to the poor is inadequate. The overall incidence of revenues is regressive.

3 Only some public expenditures for economic and social services benefit the poor more than the better off. The government has not identified individuals, groups or localities that are poor, vulnerable, or have unequal access to services and opportunities, does not have appropriate programs, and has taken only small, if any, steps to correct this. Spending on some key economic services targeted to the poor is inadequate. The overall incidence of revenues is regressive and only small steps, if any, are being taken to correct this.

4 Key public expenditures for economic and social services benefit the poor more than the better off, but some egregious regressive expenditures remain. The government has identified individuals, groups or localities that are poor, vulnerable, or have unequal access to services and opportunities, and is taking significant steps to introduce appropriate programs. With few exceptions, spending on economic services targeted to the poor is broadly adequate. The overall incidence of revenues is progressive, but some egregious regressive revenue sources remain.

5 Key public expenditures for economic and social services are well targeted to benefit the poor. There are few, if any, egregious regressive expenditures. The government has identified individuals, groups or localities that are poor, vulnerable, or have unequal access to services and opportunities, and has designed appropriate programs. Spending on economic services targeted to the poor is broadly adequate. The overall incidence of revenues is progressive, and there are few, if any, egregious regressive revenue sources.

Source: World Bank (2003b)

Table 7

2003 CPIA Score Guidelines for Transparency, Accountability and Corruption in the Public Sector

2 There are no effective audit or other checks and balances on executive power. Boundaries between the public and private sector are ill-defined, and conflicts of interest abound. Responsibilities are not clearly defined across levels of government and the reasons for and costs of decisions by public officials and the judiciary are not made clear or are not based on legal rules or procedures. Laws and policies are biased towards narrow private interests, implementation of laws and policies is distorted by corruption and resources budgeted for public services are diverted to private gain. The media are not independent of government or powerful business interests. Public officials are not sanctioned for failures in service delivery or for receiving bribes. The general public has little voice or participation in public activities.

3 Elected and other public officials often have private interests that conflict with their professional duties. Decision making is generally not transparent. External accountability mechanisms such as inspector-general, ombudsman or independent audit may exist, but have inadequate resources or authority. Restrictions on the media limit its potential for information-gathering and scrutiny, and civil society is weak.

4 External accountability mechanisms limit somewhat the degree to which special interests can divert resources or influence policy making through illicit and non-transparent means. Media publicity is an effective deterrent against unethical behavior. Risks and opportunities for corruption within the executive are reduced through adequate monitoring and reporting lines. Conflict of interest and ethics rules exist and the prospect of sanctions has some effect on the extent to which public officials shape policies to further their own private interests. Administrative corruption is low.

5 Responsibilities are clearly defined across levels of government. Accountability for decisions is ensured through a strong public service ethic reinforced by audits, inspections, and adverse publicity for performance failures. The judiciary is impartial and independent of other branches of government. The reasons for decisions, and their results and costs, are clear and communicated to the general public. Citizens can obtain government documents at nominal cost. Conflict of interest and ethics rules for public servants are observed and enforced. Top government officials are required to disclose income and assets, and are not immune from prosecution under the law for malfeasance. Authorities monitor the prevalence of corruption and implement sanctions transparently.

Source: World Bank (200b)

Table 8

Revised 2004 CPIA Categories

A. Economic Management

1. Macroeconomic Management
2. Fiscal Policy
3. Debt Policy

B. Structural Policies

4. Trade
5. Financial Sector
6. Business Regulatory Environment

C. Policies for Social Inclusion/Equity

7. Gender Equality
8. Equity of Public Resource Use
9. Building Human Resources
10. Social Protection and Labor
11. Policies and Institutions for Environmental Sustainability

D. Public Sector Management and Institutions

12. Property Rights and Rule-Based Governance
13. Quality of Budgetary and Financial Management
14. Efficiency of Revenue Mobilization
15. Quality of Public Administration
16. Transparency, Accountability, and Corruption in the Public Sector

Source: World Bank (2004a)

References

- Adam, Christopher, Gerard Chambas, Patrick Guillaumont, Sylviane Guillaumont-Jeanneney and Jan Willem Gunning. 2003. "Performance-Based Conditionality: A European Perspective." Processed.
- Burnside, Craig and Dollar, David. 2000. "Aid, Policies, and Growth." *American Economic Review*, September, 90(4), pp. 847–68.
- Collier, Paul, Patrick Guillaumont, Sylviane Guillaumont and Jan Willem Gunning. 1997. "Redesigning Conditionality." *World Development*. Vol. 25, No.9, pp. 1399-1407.
- Dalgaard, Carl-Johan and Hansen, Henrik. 2001. "On Aid, Growth and Good Policies." *Journal of Development Studies*, August, 37(6), pp. 17–41.
- Easterly, William, Ross Levine and David Roodman. 2003. "New Data, New Doubts: A Comment on Burnside and Dollar's "Aid, Policies, and Growth." *American Economic Review*.
- Guillaumont, Patrick and Chauvet, Lisa. 2001. "Aid and Performance: A Reassessment." *Journal of Development Studies*, August, 37(6), pp. 66–92.
- Guillaumont, Patrick, and Jeanneney S. Guillaumont. 1995. *Propositions pour un nouveau type de conditonalitie'*. European Commission, DG VIII.
- Hansen, Henrik and Tarp, Finn. 2000. "Aid Effectiveness Disputed." *Journal of International Development*, April, 12(3), pp. 375–98.
- Hansen, Henrik and Tarp, Finn. 2001. "Aid and Growth Regressions." *Journal of Development Economics*, 64(2), pp. 547–70.
- Kanbur, Ravi. 2000. "Aid, conditionality and debt in Africa." In F. Tarp (Ed.). 2000. *Foreign Aid and Development*. London and New York. Routledge. pp 409-422.
- Kanbur, Ravi. 2004a. "Growth, Inequality and Poverty: Some Hard Questions." <http://www.arts.cornell.edu/poverty/kanbur/GroIneqPov.pdf>
Accessed October 23, 2004.
- Kanbur, Ravi. 2004b. "What Change Does Attention to the Poorest Imply?" Paper presented to the DFID conference on Reaching the Very Poorest, London, May 12, 2004.
- Kanbur, Ravi. 2003. "The Economics of International Aid." <http://www.arts.cornell.edu/poverty/kanbur/HandbookAid.pdf>
Accessed October 23, 2004.

Tarp, F. (Ed.). 2000. *Foreign Aid and Development*. London and New York. Routledge. pp. 409-422.

World Bank. 2003a. "Allocating IDA Funds Based on Performance: Fourth Annual Report on IDA's Country Assessment and Allocation Process."

<http://siteresources.worldbank.org/IDA/Resources/PBAAR4.pdf>

Accessed October 23, 2004.

World Bank. 2003b. "Country Policy and Institutional Assessment 2003: Assessment Questionnaire." <http://siteresources.worldbank.org/IDA/Resources/CPIA2003.pdf>

Accessed October 23, 2004.

World Bank. 2003c. "2003 IDA Country Performance Ratings."

<http://siteresources.worldbank.org/IDA/Resources/QuintilesCPIA2003.pdf>

Accessed October 23, 2004.

World Bank. 2004a. "IDA's Performance-Based Allocation System: IDA Rating Disclosure and Fine-tuning the Governance Factor."

<http://siteresources.worldbank.org/IDA/Resources/PBAFINAL.pdf>

Accessed October 23, 2004.

World Bank. 2004b. *Economic Growth in the 1990s: Learning from a Decade of Policy Reform*. Washington, D.C.