

How Can Donors Create Incentives for Results and Flexibility for Fragile States? A Proposal for IDA

Alan Gelb

Abstract

International Development Association (IDA) donors and others operating a country performance-based allocation system face two difficult problems: how to strengthen incentives to produce and document development results and how to increase flexibility for fragile states. Fragile states have the greatest need for projects, but their projects tend to rate poorly in performance-based allocations systems, which provide little incentive to produce successful projects in fragile states or other countries.

This paper offers a proposal to address these issues. First, rebalance incentives to increase attention to results and to the frameworks for monitoring and evaluating them. Second, supplement performance-based allocations to fragile states through a performance fund to enable well-performing projects to be scaled up. Doing so would change the conversation about allocations to fragile states from “how much?” to “where can we show value?” and lead to a stronger and more general link between development results and development financing.



Center
for Global
Development

www.cgdev.org

Working Paper 227
October 2010

How Can Donors Create Incentives for Results and Flexibility for Fragile States? A Proposal for IDA

Alan Gelb
Center for Global Development

The author would like to thank Ginger Turner and Bruno Vincent who collaborated on an earlier study (IDA, “Selectivity and Performance: IDA’s Country Assessment and Development Effectiveness,” paper prepared for the IDA 15 Replenishment [Washington: International Development Association, 2007], on which part of this analysis is based, and also the many colleagues who have provided useful comments and suggestions.

CGD is grateful for contributions from the Norwegian Ministry of Foreign Affairs, the Swedish Ministry of Foreign Affairs, and the Australian Agency for International Development in support of this work.

Alan Gelb. 2010. “How Can Donors Create Incentives for Results and Flexibility for Fragile States? A Proposal for IDA.” CGD Working Paper 227. Washington, D.C.: Center for Global Development. <http://www.cgdev.org/content/publications/detail/1424541>

Center for Global Development
1800 Massachusetts Ave., NW
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

The Center for Global Development is an independent, nonprofit policy research organization dedicated to reducing global poverty and inequality and to making globalization work for the poor. Use and dissemination of this Working Paper is encouraged; however, reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License.

The views expressed in CGD Working Papers are those of the authors and should not be attributed to the board of directors or funders of the Center for Global Development.

Summary

International Development Association (IDA) donors and others operating a country Performance-Based Allocation (PBA) system face two difficult questions: (i) How to strengthen incentives to produce and document development results? (ii) How to increase flexibility for Fragile States (FS)? The questions are related. Needs are greatest in FS, but the success rate of projects tends to be lower in these more difficult countries, which are rated low in the PBA system. Simply increasing allocations to FS, as some have urged, therefore comes up against the concern that this should not deteriorate overall effectiveness.

Especially in FS development projects need to be seen as “venture capital”. Successes are particularly valuable given the high levels of need and the importance of showing some successes so that country programs should have flexibility to scale them up based on proven effectiveness. However, the current “performance-based” allocation system provides little incentive to produce successful projects, whether in FS or other countries. This paper offers a proposal to address these issues.

The first element is to re-balance incentives for lending staff and their counterparts, to increase attention to project-level results and to the frameworks for monitoring and evaluating them. External evaluation is now delayed to the point where any potential incentive effect is diffused. A staff-member over the age of 54 will typically retire before any assessment emerges on a project that she/he has developed. In contrast, fiduciary, social and environmental safeguards all benefit from “short feedback loops” that invite external scrutiny and bring concerns very quickly to the fore. Project design, including monitoring and evaluation (the quality of which is strongly related to project success) needs a similarly short independent feedback loop to balance out incentives.

The second element of the proposal is to supplement PBA allocations to fragile states by a performance fund, to enable well-performing projects to be scaled up. This will create incentives, both for staff working on these difficult countries and for “good ministers” to develop and implement projects that achieve demonstrable development results. It would also change the conversation around FS allocations, away from “how much” to “where can we show value”? The size of such a fund would depend on the parameters, but it might plausibly be in the range of SDR 800 million. The expected phase-down of support for post-conflict countries in IDA16 may create some space for this approach, which could be on a smaller scale if restricted to countries not receiving special allocations.

As results frameworks strengthen, the scale-up approach could be used more widely to steer funding towards better-performing projects and programs. This would tighten the link between funding and development results and encourage new output-based aid approaches as part of scaling up.

1 Introduction

As the major multilateral facility in the international aid architecture, IDA plays a central role in shaping the mechanisms of development assistance. The principles on which IDA allocates funds between countries of different types and the mechanisms and incentives used to maximize the effective use of resources are questions of wider relevance. IDA has for many years operated a “Performance Based Allocation” (PBA) system, and other donors have increasingly tended to use variants of such a system for allocating their own funds. IDA has also been at the forefront of growing concern to improve the effectiveness of aid in producing development results. This is not necessarily because its programs are considered to be less effective than others but because it is expected to be a leader in this field.

This paper considers two issues: how to increase the flexibility of IDA in dealing with Fragile States (FS) and how to strengthen incentives for the effective use of IDA resources to produce development results? These questions are related. Needs, as measured by development indicators, are greatest in the FS. In addition, their difficulties often impact on their neighbors, and this increases the urgency of supporting their reconstruction and development, and of showing some successes to build momentum. But the success rate of projects in achieving their development goals tends to be lower in countries, such as FS, which are rated low in the PBA system. IDA defines this group of countries on the basis of a low score on the Country Policy and Institutional Assessment (CPIA), the main component of the PBA. Concern to increase allocations to the FS, which are likely to become a larger share of IDA’s portfolio as more successful countries grow and graduate, therefore comes up against concerns for effectiveness:

“While IDA Deputies support exceptional assistance to fragile states, the support is mixed with potential erosion of incentives for IDA’s normal performance-based approach to aid allocation and the need to emphasize results.”
(Agbonitor *et al* 2009)

At the same time, the issue of effectiveness in FS needs to be seen in the context of increasing concern that development funds produce demonstrable results in all countries, whether in terms of income and poverty reduction, or more generally in terms of helping achieve other MDGs and critical public goods. The traditional, country-based, models of assistance are under scrutiny from many quarters and under challenge from new modalities, including sector-specific vertical funds and the resources-for-infrastructure agreements mounted by new players, including China.

Considering first the FS, we argue that development projects and programs need to be seen more as “venture capital” and provided scope for scaling up successes. These are particularly valuable in such countries, given the high levels of need. Such an approach requires a modified allocation process, one that can scale up country allocations in response to very strong project performance. Without such a demand-led

element, there is the risk that simply pushing more money towards FS, as some have urged, will deteriorate overall effectiveness.

At the same time, implementing such a system will require increasing attention to project-level results and to the frameworks and processes used to monitor them. This brings us to our second topic, the incentives for a strong results focus. IDA has introduced a results monitoring framework which provides more information on project deliveries than was previously available. But the current system for allocating funds to countries provides negligible incentives to country programs staff and their country counterparts to produce successful projects. External evaluation feedback in the area of results is delayed, to the point where staffing and counterpart turnover diffuses any potential incentive effect. Indeed, incentives in this area face increased competition from fiduciary, social and environmental safeguards. The latter all benefit from “short feedback loops” that invite external scrutiny by interested parties and bring concerns very quickly to the notice of staff and management. Certainly, projects should “do no harm” -- but they should also “do good”. Even though some results may take a long time to materialize, project design and implementation would benefit from a similarly short feedback loop. As results frameworks strengthen, the scale-up approach could be used more generally to steer funding towards better-performing projects and programs, and to further strengthen the links between IDA funding and development results.

Section 2 considers the current PBA system from the perspective of these two problems. The system has been critiqued on many grounds, but it has a number of strong points, including an element of insulation against politically-driven allocations. It also has the ability to direct funds towards countries that are better-performing in terms of overall development outcomes, as well as having generally better policies and institutions, at least as assessed by the CPIA ratings that underpin the PBA. However, despite the increased concern to demonstrate direct results from development funding, the system provides negligible direct incentive for strong performance at project level. Moreover, its relatively rigid approach towards FS has needed to be modified through a series of *ad hoc* exceptions and special allocations. Some observers have urged changes in the formula to channel an increasing share of funding to such developmentally challenged countries and to eliminate exceptional treatment.

Section 3 considers indicators of development effectiveness at project level and the incentives to measure and monitor them accurately. This raises a number of complex questions: What should be considered under “effectiveness”? How to account for the less tangible benefits of development support? Does an increased focus on measurement runs a risk of driving out developmentally meaningful or “transformative” objectives in favor of narrower, more specific but less developmentally meaningful indicators? A full treatment of these issues is beyond the scope of this paper but they lead to consideration of the monitoring and evaluation frameworks used to evaluate effectiveness. Projects with strong monitoring and evaluation frameworks have higher rates of success, but a recent by the World Bank’s Independent Evalua-

tion Group (IEG) concludes that they are weak in most projects. As noted above, one reason for this may be incentives -- the far slower external feedback loop on project effectiveness and results relative to feedback loops on social, environmental and fiduciary safeguards which compete for the attention of project managers.

Section 4 offers proposals for change in two areas. First, modify the PBA approach for FS by providing an additional project-performance-based tranche to these countries, based on demonstrated capacity to implement projects effectively and scale them up. This would require additional efforts to strengthen performance assessment, including the monitoring and evaluation frameworks for the projects, and to ensure timely decision-making, but these are manageable. Second, introduce a short-term advisory feedback loop with external input that focuses on a project's ability to monitor and measure potentially developmentally meaningful outcomes and impacts. While a number of institutional arrangements are possible, it is important that this loop have a degree of independence and that its conclusions are made fully public.

Neither of these suggestions is intended to replace efforts to address the problem of results at the level of the aid modality itself, including through new Results-Based Instruments or "COD Aid" (Birdsall and Savedoff 2010) which would link disbursements to results directly at project level. The proposals are intended to provide additional incentives to move in such directions. The more such instruments can be developed, the easier it will be to move towards strong results frameworks and genuinely results-based allocations.

2 Debates on "Performance-Based Allocation"

While there have been a number of modifications to IDA's performance-based allocation system, the essentials have been maintained for some time.¹ The standard per capita allocation to a country starts off with the Country Performance Rating (CPR), which depends mainly on the CPIA, an annual assessment of policies and institutions. Cluster "D" of the CPIA, which measures performance in governance-related areas, is super-weighted in the formula, and accounts for 68% of the weighting. The other three clusters, macroeconomic management, structural policies and social and related policies, together account for 24% of the weight. Portfolio performance assessments made annually by the region responsible for the country program are included in the CPR but only with a weight of 8%.² The total allocation formula multiplies the fifth power of the CPR with population and with GNI/head raised to the power of -0.125 to provide more IDA to poorer countries. The resulting allocation has a high degree of selectivity and produces considerable

¹ The PBA system is outlined in IDA 2010. For more detail and some proposed alternatives, see IEG 2009c, Leo 2010, Guillaumont *et al* 2010 and Alexander 2010.

² The Portfolio Performance Rating (PPR) is measured by the percentage of problem projects in the portfolio: see IDA 2010 Annex 3.

dispersion in the IDA allocation. A country rated high on the CPR can receive 6 or 7 times the per-capita allocation of a low rated one.³

Many countries do not actually receive the standard PBA allocation. Very small countries benefit from minimum allocations, while allocations are capped for some very large countries, in particular India. Special windows provide post-conflict and re-engagement funding to about half of the countries rated at below 3.2 on the CPIA scale, the rating below which countries are considered as Fragile States. Although the exceptions are intended to be transitional, the post-conflict allocation was extended from 5 years to 10 years in IDA 15, and the re-engagement window from 3 to 5 years. A regional window provides funding to supplement country allocations for common projects that involve three or more countries, such as regional infrastructure. Finally, since the grant component of an IDA credit is about two-thirds, countries that receive grants because of concern over debt sustainability have a downward adjustment in the program amount.

This approach and the resulting allocations have been criticized from a number of perspectives. Some have focused on the CPIA, characterizing it as a “one-size fits all” approach to development policy, driven by Washington Consensus-style thinking that leaves little policy space for aid recipients. Partly for this reason, some have argued for more direct weighting on the actual development performance of a country, as measured by country-level development outcomes rather than policies (Kanbur 2005). Another body of studies argues that the need for *ad hoc* provisions demonstrates that the basic formula is unsatisfactory, particularly in dealing with fragile and disadvantaged states (Guillaumont *et al* 2010). They urge the elimination of most of the *ad hoc* provisions and the inclusion of indicators of Economic Vulnerability and (low) Human Assets which make development more difficult. The effect of including these would be to shift more funding towards the FS. However, some that now receive special allocations would lose out with the proposed changes.⁴

A detailed assessment of these views is beyond the scope of this paper, but a few points could be noted. The CPIA has gradually evolved over time; following review by an external panel in 2004 and subsequent revisions it now covers 16 distinct areas, down from a previous 20.⁵ Close examination refutes the argument that it simply reflects a one-dimensional Washington Consensus view of “good” versus “bad” policies. Only a few questions can be interpreted directly in this way; the rest

³ Similar systems are used by the regional development banks to allocate concessional resources. Although some bilateral donors do use indicator-driven rating systems to allocate resources, these are not usually explicit and public; the Millennium Challenge Corporation is a notable exception. Bilateral systems are essentially advisory since they are subject to political override.

⁴ Although their share has fluctuated, on average FS have received in total some 10%- 15% of IDA allocations in recent years. In IDA 15, IDA allocations were estimated at SDR 2.5 billion, about 10% of total resources. Most of these were to post-conflict and re-engaging countries which benefited from special allocations. (IDA 2010).

⁵ For details of the assessment and areas of revision see IDA 2005.

require pragmatic expert assessments of the quality of various country systems.⁶ In many cases these will be impossible to make without an understanding of the outcomes produced or services delivered by these systems. In this sense, it can be argued, the CPIA ratings may be at least partly endogenous to country development results (IDA 2007).

Nevertheless, the PBA formula has a high degree of arbitrariness, and it is difficult to substantiate specific features on empirical grounds. The uniform weighting of the components in the total CPIA, for example, is not based on theoretical or empirical justification. Even though much recent research points to the importance of institutions (Acemoglu, Johnson and Robinson 2004), there is no empirical basis for the particularly high weighting on the governance component in computing the CPR⁷. It is also hard on equity grounds to justify the capping of a large country such as India, even though it may contain more poor than all of Sub-Saharan Africa (SSA).⁸ This capping has an important effect; it doubles IDA flows to SSA (and also to some small countries outside Africa) relative to what their level would be in the absence of the caps.

Empirical research also does not provide a very firm foundation for any particular country-based allocation formula. The concept of allocating aid based on poverty and policy is rooted in two propositions: (i) aid is more effective in better-managed countries in terms of producing country-level development results, and (ii) the marginal product of aid declines as its level increases. These propositions are reasonable, appealing, and very likely true in some sense. Project-level data (below) provide at least suggestive support for the first, and even proponents of a “big push” will agree that beyond a certain point the marginal product of aid to a country is likely to decline. However, the fragility of cross-country results relating aid and policies to development impact shows the difficulty of using cross-country estimates to define an optimal poverty-policy frontier for a given level of aid.⁹ Moreover, as observed by Barder 2010, even if such an optimal allocation formula were proved for aid in gen-

⁶ The influence of the Washington Consensus is most evident in questions 4 (trade policy) and 6 (business regulation). Trade policy ratings are based on a 75% weight on trade restrictiveness and 25% weight on the quality and effectiveness of customs administration. A rating of 4 on the former requires an average tariff of no more than 16% and a maximum of 25% or less, 4 or fewer tariff bands, sparing use of non-tariff barriers and no export taxes. Other critiques of the CPIA argue that some criteria value efficiency over equity; for example, Alexander 2010 notes the emphasis on low-distortion taxes such as VAT relative to effective enforcement of income taxation on high-income taxpayers.

⁷ Statistically, cluster D of the CPIA is highly correlated with the other three clusters so that it is difficult to determine any independent effect of governance on performance within the CPIA, even though the weighting will, of course, matter for allocations to individual countries.

⁸ Using the Multidimensional Poverty Index Alkire and Santos 2010 estimate that India contains 645 million poor relative to 458 million in Sub-Saharan Africa.

⁹ There is still active debate on whether aid has any robust impact on development results at all, once policies and country-specific factors are taken into account (Rajan and Subramanian 2008, Arnt, Jones and Tarp 2009), although a consensus may be developing that it has a small positive effect.

eral, the decision for any particular source, such as IDA, would have to depend on the allocation decisions of all other donors.

Nevertheless, the PBA allocation process has a number of attractive features (IDA 2007). Development progress over 10 year periods, as measured, for example, by increases in the Human Development Index (HDI) or decreases in infant and under-5 mortality rates, is faster for countries with higher CPIA scores. Progress is faster for countries with initially low scores on the development indicator (conditional convergence), for countries not impacted by high rates of HIV/AIDS, and for countries not located in Africa. The latter progress at about twice the rate for African countries, given the initial value of the development indicator, policy and HIV/AIDS incidence¹⁰.

These results do not of course show that the countries with higher CPIA scores progress more rapidly *because* they receive higher allocations of IDA/head. Since standard IDA allocations and CPIA ratings are close to collinear it is not possible to distinguish the relative contributions of IDA flows and policies, as measured by the CPIA. Nevertheless, they indicate that the PBA system works, in the sense that it enables multilateral funders to “place their bets on the horses that run faster”. This feature does not, of course, prove impact, but is at least suggestive and valuable in advocating for IDA. It is not found for bilateral aid.¹¹

In addition, the negative Africa dummy provides a caution to those arguing for a simple and direct link between outcome performance at country level and aid allocation. Even if development outcomes were widely available at country level with sufficient timeliness and accuracy to shape allocations, they may improve less rapidly in some countries for a number of exogenous reasons. These include being landlocked and far from markets and the sea, or in a bad neighborhood. If a country in conflict reduces the growth rate of its unfortunate neighbors by 0.5% as well as its own growth by 2% (World Bank 2003), one with several borders inflicts potentially as much damage to growth externally as inside. This situation of “dynamic disadvantage (Wood 2008) will be more likely in Africa, with many small countries and borders, and a high incidence of conflict-affected states. The negative Africa coefficient can therefore be seen as an argument for capping allocations to large countries outside Africa to divert more IDA to Africa, because location in Africa can be seen as an indicator of dynamic disadvantage in achieving the MDGs.

Given these points, how then should we interpret the allocation process for IDA and the similar approaches in other multilateral development banks? These formulae

¹⁰ Since we have argued that the CPIA could be at least somewhat endogenous to development outcomes, the conclusions above should not be interpreted as simply asserting that a high CPIA causes better development outcomes. It is rather that the rating reflects the view of the analyst on the quality of the policies and institutions that are considered to be underneath the outcomes. Consistent with this view IDA2007 finds that ratings are not affected by outcome changes that plausibly reflect external shocks.

¹¹ IDA 2009.

have to be seen as the outcome of political compromise. They reflect a balance of donors' interests as between a number of characteristics: country need, as indicated by low income level; policy and especially governance; preferences for particular geographic regions over others (some donors have less interest in India, for example, than in Africa where they have many ex-colonies); small states; and perhaps to a degree regional public goods such as infrastructure or regional bads such as conflicts which are liable to spill over onto neighboring countries. They also provide a way for donors to signal support for countries considered to be improving economic management in general and governance in particular.

Because of the status of the formula as an outcome of political compromise, and also because part of the benefit of multilateral aid is expected to be its predictability, suggestions for major changes to need to be weighed carefully. Even those critical of the current system recognize the value of a rules-based allocation system relative to one driven by political considerations. The approach is also practical -- the country and portfolio ratings are made on an annual basis, whereas deeper development indicators, such as poverty or mortality levels, may not be available on a timely basis for all countries.¹²

Nevertheless, from the present perspective the system has two limitations. First, the smallness of allocations to countries rated low on the PBA scale leaves little flexibility to take advantage of windows of opportunity, except through major categories of exceptions which can themselves be targets of political pressure. Second, even if the process is successful in allocating funds to countries with generally stronger institutions and development outcomes, the allocation process itself provides very little incentive to work towards more effective use of development funds.

To illustrate this point, consider the effect of a 10% increase in the CPIA (uniform across clusters) versus that of a 10% increase in the portfolio rating, each starting off from middling baseline ratings of 3.5. The first increase yields a gain of 55.3% in the IDA allocation. The second only produces a 4.1% gain. This is far smaller even than the 35% change in the allocation that would correspond to an estimate of one standard-deviation of a CPIA rating.¹³ Incentives that work through the allocation system are therefore far stronger in the direction of contesting the CPIA rating than in the direction of improving the portfolio.

3 Development Effectiveness and the Quality of Project Ratings

¹² The European Union includes some results-based tranches in their country-level budget-support programs (European Commission 2005). This may be a trend to consider for the future, provided expected results can be calibrated to country circumstances.

¹³ Gelb, Ngo and Ye 2004 use differences between the World Bank and African Development Bank scores on very comparable CPIAs to derive an estimate of 0.24 for the standard error of a CPIA rating. The two sets of scores were not public and were not shared for the years compared, so they can reasonably be considered as independent efforts to assess policies and institutions. The effect of the 10% improvement in the portfolio rating would be offset by a measurement error in the CPIA of only 0.05, far smaller than the estimated range of precision.

Especially since the fall of the Soviet Union diminished the strategic arguments for assistance, IDA and other aid programs have been subject to increasing pressure to demonstrate development results. One response is the Results Management System (RMS) for IDA, launched in 2002 and strengthened in 2005.¹⁴ While this allows for more specific monitoring, it leaves open three further questions. How to approach the Fragile States? How adequate is the quality of the monitoring and evaluation (M&E) systems on which project ratings are based? Does feedback from independent evaluation of project outcomes provide an incentive to design more successful projects, as well as opportunities to learn?

Fragile States and Project Outcomes. Countries rated low on the CPIA scale pose a particular challenge of need versus performance relative to other poor, but well-governed, countries. Fragile states are defined as those with scores below the level of 3.2. Development needs, as measured by children out of school or mortality rates, are typically much higher in such states; often they are twice as high (World Bank 2007). While individual countries do transition out of this group (Uganda, Mozambique) and into it (Zimbabwe), membership shows a high persistence over time. In 2000-2008 the median number of FS was 33 and the number of chronically fragile states was 28 (Agbonitor *et al* 2009). Of the 33 countries, 20 were in Sub-Saharan Africa with Nigeria the only country exiting in the last 10 years (Ramsey 2009). As other developing countries grow and continue to make progress, the differentiation within the low-income group between those able to move forward and those stuck for long periods in situations of poor governance and policies and slow, or no, development is only likely to become starker.

Some argue that the citizens of Fragile States are in effect “penalized twice”: first by poor governance, and second by low aid allocations provided by IDA-type PBA systems that place a heavy weight on governance (Guillaumont *et al* 2010). The limitations of this argument emerge from data on project implementation. The most comprehensive set of data on project performance comes from the Independent Evaluation Group of the World Bank (IEG), which determines three ratings, usually between one and three years after project completion: success in meeting project outcomes (scale 0-6), sustainability (0-4) and institutional impact (0-4). These ratings are correlated, and we consider therefore only project outcome ratings.

Data from a sample of 4,370 projects evaluated by IEG over 1980 – 2004 show that the probability of project outcomes being assessed as satisfactory is strongly associated with the CPIA rating of the country, whether at the start of the project, at the end, or on average during the implementation period (IDA 2007). Table 1 shows the

¹⁴ The RMS provides indicators at two levels. At country level, fourteen development indicators are tracked. To measure IDA’s contribution, the RMS tracks: (i) country strategy (number and clarity of results-based Country Assistance Strategies); (ii) project outcomes (including measures of quality at entry, IEG assessments of project outcomes and the share of project reports with an adequate baseline to monitor progress); (iii) selected project outputs in four key sectors and (iv) statistical capacity in IDA countries.

estimated probability of given project outcome ratings for two typical countries with average CPIA ratings over the project implementation period of 2.5 and 4.5 respectively. More than two thirds of the projects in the high-rated country are rated either fully or highly satisfactory (4 or 5), but only one quarter are rated at these levels in the low-rated country. Conversely, 56% of projects in the low-rated country are rated as highly unsatisfactory, unsatisfactory or partly unsatisfactory (0, 1 or 2), but only 19% in the high-rated country. The proportion of partly satisfactory projects (3) is similar. A one point increase in the average CPIA about doubles the probability that the project will be assessed in the top rating.

Table 1
Probability of IEG Project Ratings by CPIA
(percentage of total projects)

IEG Rating	CPIA = 2.0	CPIA = 4.5
0	5	1
1	45	14
2	6	4
3	19	15
4	24	61
5	1	5

Source: IDA 2007. Probabilities from ordered logit.

Recent assessments using later data confirm this picture, while noting that there may have been some improvement in later years, with project performance in low-rated countries somewhat closing the gap as part of a general trend towards improved project outcome ratings. A recent assessment of the Fragile State portfolio in IDA programs reiterates the greater difficulty of achieving strong outcomes in these countries (Agbonitor *et al* 2009). Even though, as discussed further below, the project ratings system is based on information that falls short of a rigorous analysis of project impact, the ratings suggest that IDA is indeed more likely to be effective in high-rated countries. One third of projects in the FS are at risk, more than double the rate in other IDA countries. Simply increasing the share of IDA allocated to low-rated countries is likely to deteriorate overall portfolio effectiveness.

Could these results be spurious? Are the CPIA ratings simply reflecting the project ratings? The average implementation period of projects is about 6 years¹⁵, to which must be added the interval between project completion and the initiation and completion of the IEG evaluation of the project completion report. With this very long feedback cycle of 7-8 years the CPIA ratings precede the IEG ratings by several years, so that it is not likely that the project ratings are able to influence the country ratings. Are the project ratings influenced by the then-current CPIA ratings rather than actual estimates of project success? This too is not plausible. The IEG rating

¹⁵ This average excludes Development Program Loans (DPLs) or Poverty Reduction Support Credits (PRSCs) which disburse more rapidly.

system follows an objectives-based approach, where success is judged against stated project objectives and not on country-wide policies or outcomes. Their ratings may also diverge from the final project ratings of the task manager, usually in a negative direction (referred to as the “disconnect”: see below).¹⁶ These factors suggest that reverse causality from project rating to country ratings is not likely to be a problem, and that the project ratings do have independence from country ratings.

While country conditions affect project performance they do not determine it fully. Ramsey 2009 analyses the determinants of project performance in the FS of Sub-Saharan Africa. Project implementation support and supervision makes a difference, particularly in the first two years after Board approval. Successful and failing projects differ in the frequency of supervision missions, in their intensity (intense missions are defined as those that include senior staff other than the team leader and the fiduciary team), and also in spending on supervision over the first two years (Table 2). This confirms the importance of incentives, both in terms of senior staff and budget, as well as sustained attention.¹⁷

Table 2
Indicators for Successful and Failing Projects in African FS
(percentage of projects and \$ 000)

	Successful	Failing
High Frequency of Supervisory Missions (%)	50	33
High Intensity of Supervisory Missions (%)	83	0
Implementation and Supervision Support in First two years (\$ 000)	153	108

Source: Ramsey 2009.

Some have argued that differences in country ratings should mainly affect the modalities through which aid is delivered rather than the total levels of aid (Guillamont 2010). Indeed, this is the practice; certain instruments, such as budget support tend to be offered more frequently to better-performing countries with stronger institutions, especially in the area of budget and financial management and especially in SSA (World Bank 2004, IEG 2009a). There may be more scope for differentiating aid modalities in the future. However, the results above already take into account some degree of endogenous portfolio adjustment to country conditions.

The difference between project ratings in high and low CPIA countries suggests that development finance needs to be looked at differently between the two sets of countries. In the high-rated countries, there is a reasonable presumption that aid will be

¹⁶ Most IEG evaluations are desk reviews of project documentation. About 20% of projects are selected for more in-depth review, including a field visit, at a cost of approximately \$25,000 per project.

¹⁷ Ramsey 2009 argues that projects in non-fragile states can often overcome supervision and support weaknesses in the first two years but that this is not true for projects in FS. Her results can be interpreted in different ways; for example the zero incidence of high-intensity supervision could also imply that senior staff visit projects on a selective basis.

well-used, at least as assessed by current methods. These countries are also typically more stable in terms of their CPIA ratings, and this argues for a longer-run commitment in terms of their country aid envelope. Eifert and Gelb 2008 show that multi-year commitments are possible for such countries with only small risk, provided there is the possibility of re-contracting in the case of a major change in governance and policies. Country-level allocations can be expected to work reasonably well in these conditions, at least until more direct measurements of effectiveness and impact of aid can be institutionalized.

In low-rated countries the situation is different. The probability that a given project will fail to achieve its development objectives is quite high; indeed, excessive concern that every project achieve all of its objectives is likely to lead to over-cautious projects with limited outcome expectations. In these countries, development finance should be thought of more as venture capital. The expected rate of success may be lower but the payoff to successful development efforts will be particularly great, given the extreme needs. Such projects and programs need to be well-resourced, encouraged, expanded, held up as examples and recognized in staff performance assessments.

M&E Systems and Project Outcomes. Whether projects should be evaluated against specified objectives, as is currently the practice for IBRD and IDA projects, or against norms and standards (such as an economic rate of return (ERR) or other criteria), as is more the practice for IFC and MIGA projects, is a debatable question. Relatively simple projects focused on particular investments can find the latter more appropriate, but it is less easy to do convincingly for complex projects which include policy and institutional change objectives as well as more specific physical investments. Especially for such projects there are many unknowns along the chain running from project inputs to outputs and then to outcomes. It is therefore not always easy to estimate a convincing ERR, although the effort to do so will force attention to the question of what the project is trying to achieve.¹⁸

Nevertheless, only including the immediate outputs raises the risk that the project's objectives lack relevance to development objectives. For example, support to a road fund aimed at improving road maintenance and the percentage of roads in good condition will be of relatively little use in reducing the costs of transport if the real problems are corruption and delays at ports and customs, unofficial roadblocks and the cartelization of road transport activities.¹⁹ For this reason, independent assess-

¹⁸ The Millennium Challenge Corporation does require projects to provide an ERR as well as specify outputs, probably in part for this reason.

¹⁹ Marteau and Rabilland 2007 note that the condition of roads may not be the main reason for inefficient and costly transport. Often the most important problem lie in ports where delay times are twice those of road travel. Inefficient port processes in Douala, Cameroon, contributed to the delays and high cost of transporting goods to N'Djaména in Chad, 2,000 km from the sea. Other problems found in their study included border delays, cartels in the trucking industry, multiple clearance processes, and bribe-taking, all of which kept transport costs artificially high. The five-week journey over rail and road required seven docu-

ments by the IEG include three dimensions: efficacy (were the objectives achieved?); efficiency (were they achieved in a cost-effective manner?) and relevance (were the particular objectives relevant to some plausible and valued development goal?).

Considering the complexity of these questions and the difficulty of measuring certain important institutional goals such as building capacity in a particular area, there can always be room for disagreement and for unrealistic expectations of what can be measured.²⁰ But some level of measurement is essential. If development programs are to be meaningful, there will be a fuzzy frontier of good practice and some clear indications of initial design flaws likely to prevent a project from demonstrating results. These include the absence of a baseline and benchmarks for comparison, the absence of (independent) mechanisms for monitoring progress against these expectations, or even the absence of any monitorable contributions at all.²¹

A recent IEG assessment of the monitoring and evaluation framework for 748 projects assessed over the period throws some light on these questions. M&E frameworks were classified into four categories: high, substantial, modest and negligible. Table 3 shows the distribution of these ratings, both overall and relative to the distribution of the project ratings on development effectiveness. The first conclusion is the low average quality level of the M&E ratings; most systems are rated as either negligible or modest. Very few are rated as high quality; despite increasing attention to this area there is still a long way to go. This underpins the assessment of IEG 2009b which notes that although: “*The Bank has offered general encouragement, detailed guidelines and specific requirements aimed at this goal [strong M&E systems] for over a decade...*” (p25)... “*The low M&E ratings...suggest a continuing need to revisit the incentives, use and resourcing of M&E in Bank-supported projects.*” (pxiv).

ments and suffered from poor and fragmented trucking services, widespread rent-seeking resulting in many checkpoints, security problems, and weak customs administration in Chad.

²⁰ There is also a risk that an assessment carried out at the end of a project with many more years of information can ‘second-guess’ judgments made at the concept stage of the project. Some projects may be lightly or more formally restructured in response to changing conditions or information, and this raises the question of what set of objectives to evaluate them on.

²¹ Natsios 2010 argues against an excessive focus on measuring impact, on the grounds that this forces a retreat into immediate outputs of lesser developmental relevance at the expense of more transformational programs with stronger institutional goals. While this can be a valid concern, a credible program will still need some convincing form of assessment, at least for the bulk of its activities. Lack of credibility also opens up a program to his worst-case nightmare scenario -- ever-more intrusive process and input controls from the “counter-bureaucracy”. Moreover, institutional change is not impossible to measure. For example, the Public Expenditure and Financial Accountability (PEFA) indicators provide a well-established metric for public expenditure and financial management; other benchmarking systems have been developed for areas such as procurement and logistics performance. Even in very difficult areas of institutional change innovative studies have been able to test for impact; for an example for US support for democracy see Finkel *et al* 2006.

Table 3
Project Outcome Ratings by M&E Quality Rating
(Number of projects)

Project Rating	M&E Quality Rating				Total
	Negligible	Modest	Substantial	High	
Highly Unsatisfactory	1	0	1	0	2
Unsatisfactory	29	23	1	0	53
Marginally Unsatisfactory	36	53	13	0	102
Marginally Satisfactory	52	165	49	1	267
Satisfactory	4	108	143	15	270
Highly Satisfactory	0	4	17	5	26
Total	122	353	224	21	720

Source IEG Data.

The second feature is the relationship between the M&E and project outcome ratings. In principle this could go either way, since strong monitoring and evaluation could uncover problems that cause results to be downgraded. This is not the dominant story here. While there are many outliers, including projects that are surprisingly well-rated on outcomes despite a low M&E rating, projects with a low rating on M&E are far less likely to be assessed highly on achieving their development results.²² Three quarters of the projects with an M&E rating of Substantial or better are rated as Satisfactory or better, compared with only one quarter of those with M&E rated as Modest or less. Whether this reflects the better planning and preparation associated with a strong M&E framework, or simply the limited evaluability of projects with poor M&E is not clear. Plausibly it is a combination of these effects.²³ In any event, the ratings suggest that the current estimates of project impact are based, in many cases, on less-than-adequate information on what they actually achieve.

Incentives and the Feedback Loop. While projects are rated continuously during their implementation, an analysis of the profiles of these internal ratings suggests a lack of realism. Longstanding problems are often surfaced only near the end of a project, and even then the ratings of the project team are frequently higher than

²² Since the M&E ratings and the project success ratings are made as part of the same assessment, even if they measure different things they cannot be considered as fully independent. The cases of low M&E ratings and high project ratings does raise the question of the basis on which the latter are made. In some cases these are in middle income countries, suggesting that other sources of information, such as country data, are being drawn on to assess performance. The relationship between M&E and project ratings is not simply due to country policies and institutions as estimated by the CPIA.

²³ Other studies have found a link between the level and quality of economic and sector analysis (ESW) preceding a project and its IEG development outcome ratings; still others a link between the use of research in ESW and the assessed quality of the ESW. This suggests that any impact of a sound M&E framework is likely to reflect the benefits of clearer up-front thinking as well as a measurement effect.

those made slightly later by external assessments.²⁴ But external evaluations of project performance and M&E systems come very late in the project cycle. The long feedback loop, on average some 7-8 years from project preparation to feedback, makes it difficult for the evaluations to contribute to real-time learning. During such a period project staffing will typically have changed – sometimes several times. Country counterparts and regional and country program management will also have turned over. Given mandatory retirement for IDA staff at age 62, staff over the age of 54 involved in the start of a typical project will probably be retiring by the time the project is evaluated. Staff over the age of 50 will probably be in their final assignment, with no further prospect of career development.²⁵ As in the ill-fated US housing finance system, the originators of the loan will rarely be in place when subsequent problems surface to become public, and strong incentives to originate loans will likely result in less attention to quality.

In contrast, social and environmental safeguards and fiduciary safeguards have very short independent feedback loops which are likely to influence the same staff who prepare and approve an operation. The Inspection Panel feedback loop for example, begins when it receives a request for inspection claiming that the Bank has violated its policies and procedures. This can come as early as a year or two after the project is approved or even – as in the case of the South Africa Eskom support project – before Board approval.²⁶ Once the Panel has registered a Request for Inspection, the Bank’s management has 21 days to provide the Panel with evidence that it complied with, or intended to comply with, the Bank’s relevant policies and procedures. After receiving management’s response the Panel has 21 days to determine the eligibility of the Request. Once it has been determined that the eligibility criteria have been met, the Panel may make a recommendation to investigate. If the Board of Directors approves, the investigation phase may take a few months. The total feedback loop can thus involve a lag of only some 6 months.

It would therefore not be too surprising if project preparation and Regional quality reviews paid close attention to social, environmental and fiduciary safeguards and less to indicators of project effectiveness and the quality of the monitoring and evaluation system. There is at least anecdotal evidence to support this.

²⁴ For details see IEG 2009b. IEG has typically rated the percentage of satisfactory projects about 15% lower than the Bank’s self-assessment in its final Implementation Status and Results Report made just before the completion of implementation. The difference in ratings is termed the “net disconnect”.

²⁵ Development Policy Loans have a shorter feedback loop as they disburse rapidly.

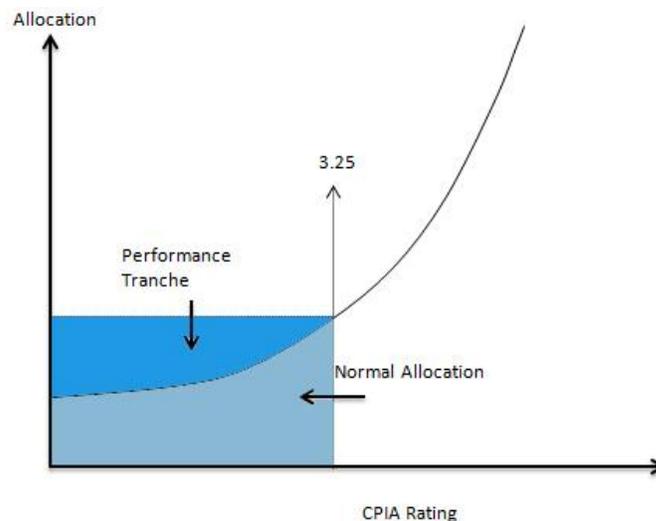
²⁶ An Inspection Panel request regarding the Eskom project was submitted on April 6, 2010. The loan was approved on April 9, three days later.

4. Towards Incentives for Results and Flexibility for Fragile States.

4a. Flexibility for Fragile States.

The first proposal, to recognize the “venture capital” side of working in FS, is to supplement the standard PBA-based allocation by a Performance Tranche to enable the scale-up of successful projects. This tranche would increase the potential allocation up to an upper bound, for example, equivalent to the financing which the country would receive at a higher “threshold” CPIA or CPR rating. For Fragile States, one possibility could be to set a threshold corresponding to a rating of about 3.2 (see Figure 1)²⁷. Access to the Performance Tranche would require the country program to demonstrate projects or programs with strong M&E frameworks that were being implemented fully successfully according to the IEG rating scale, and capable of being scaled up. A full portfolio approach that required all projects to show good results would not be consistent with the venture-capital model of aid in FS which needs to accept that some will very probably not succeed and that the goal is to ensure that successes can be scaled up. The choice of which projects to propose for this scale-up, or whether to propose any projects at all, would be that of the country program.

Figure 1
Allocation and Performance Tranche



How would such an arrangement work? It would need to address three issues: the timeliness and credibility of the project assessments, the funding of the program, and the implication of less predictable financing flows at country level.

²⁷ This particular threshold would of course provide little incentive for a country already close to the 3.2 mark. One can imagine other possibilities, such as a uniform percentage tranche for all FSs.

Timeliness and Credibility. Operational performance reviews currently carried out in real-time would be strengthened in two ways:

- “Real-time” evaluation would need to be based on a robust results framework that is monitored and evaluated continuously. This would need to include an element of independent feedback, including from local evaluation of the project in real time by civil society. Civil society in this context does not have to be limited to the beneficiaries although they should be included; it could range from professional associations to parent-teachers associations to local community NGOs to a professional independent think-tank empowered to assess project performance. This M&E system will require baseline data, which is now an IDA requirement for all projects, and therefore should be built into the design of the project.
- Independent validation. Because of the tendency for project ratings made by Operations to be optimistic relative to those made independently there would need to be a process to co-validate proposals for increased funding in a timely way, to include a degree of independent review.

Funding the Program. Funding for the program would need to cover both the additional resources allocated to the countries and any incremental resources required for implementing the augmented program.

- IDA 15 allocations to FS will be about SDR 2.5 billion or 10% of the total. The prospect for IDA 16 is a reduction to 6.7%, as exceptional allocations wind down and countries return to their normal PBA allocations (IDA 2010). This might open space for a Performance Tranche of up to SDR 800 million that would introduce flexibility and provide incentives for performance. It could partly replace the current special allocations as countries transition to normal PBA status.

The cost of implementing such a program for the FS that are not in special regimes would be more modest. Consider, for example, the 14 countries with 2008 CPIA less than 3.25 and not included in post-conflict or re-engagement special allocations. Their annual IDA allocation over the IDA15 cycle is about SDR 170 million²⁸. With a CPIA of 3.25 their total allocation would be about \$236 million, an increase of SDR 66 million per year or SDR 198 million over the three year IDA cycle. While this represents a possible 52% increase in their funding, it is a small percentage (0.7%) of the total of IDA.

A preferable option would be to have this amount additionally available through a multi-donor “scale-up” trust fund, so that allocations to better-

²⁸ Cameroon, Chad, Comoros, Djibouti, The Gambia, Guinea, Guinea-Bissau, Kiribati, Sao Tome and Principe, Sierra Leone, Solomon Islands, Tajikistan, Tonga and Yemen. These estimates are approximate.

performing countries could be sustained without reducing the IDA to better-performing countries.

- Similarly, the costs of implementing the supplemental funding should be modest, and accommodated by shifting budget resources in line with IDA flows. As the Bank now implements many projects partly funded by other donors, the use of trust fund resources would not raise any new issues. Based on the current cost of field-based IEG evaluations of complete projects of about \$25,000, the costs of mounting the external assessments would be modest.

Predictability of funding: While the introduction of a performance-based tranche would in some sense reduce the country-level predictability of the funding program, this is less an issue for low-rated countries whose CPIA scores are in any event more volatile than those for high-rated countries. Financing needs to accommodate to country conditions in these countries, whether “turnarounds” or relapses; long-run commitments are more risky. At the project level, there could be a time lapse of several months between an assessment and the release of supplementary funds, but since the average project implementation period is some 6 years (with a mid-term review after 3 years) this should not be a major problem.

4b Incentives for Results.

The second proposal is to dramatically shorten the independent feedback loop on the evaluability of projects and the quality of their development outcomes. This could involve building in an advisory review of these elements of the projects at concept stage, with independent input. Reviews would point out strengths and weaknesses, and provide a clear indication of whether the project, and its associated data, could be expected to be evaluable.²⁹ They would also be posted on the web. This leaves all options open to management, including that of proceeding with projects lacking sound M&E systems, but it would create a strong incentive to improve the results focus and evaluability of projects from the start. It would also provide clearer feedback to task managers, who currently cannot benefit from a sterling IEG evaluation of project design because this comes so long after the fact. In this way it would also spur efforts to develop innovative, results-based, projects.³⁰

²⁹ Unlike current IEG review and the reviews undertaken by the now-defunct Quality Assurance Group (QAG) these reviews would not necessarily emphasize formal ratings. They would, however, draw clear attention to issues relating to design and to potential evaluability.

³⁰ Some steps in this direction have recently been taken by the Bank, including the creation of internal Bank-wide panels to vet project and M&E design. The new policy on information disclosure will also help to enable external scrutiny of project design and implementation. These are steps forward, but would provide less focused and independent external feedback on project design than the approach proposed. Whether such an upstream role should be performed by the IEG itself is an open question; there can be concern that moving upstream in this way would compromise a later independent role in project evaluation. However, it would certainly be appropriate for IEG to set out standards and approach for this

5. Conclusion

If IDA has only to show that it can channel resources to countries with stronger countrywide development outcomes, the present PBA system is probably adequate. But to remain a central point for the development architecture, it will have to demonstrate two additional things. First, on-the-ground relevance in Fragile States, which will become a larger share of its portfolio as better-performing countries graduate. The proposal above would re-position the debate on the allocation of resources, towards a framework that recognizes important differences between high-performing and low-performing countries. Pressure to allocate more IDA to low-performing countries should confront the special need to ensure that resources are well used.

This would change the conversation around FS, from “how much IDA” to “where can we really add value”; also from commitment levels to results. No low-performing country would be denied resources above the current standard allocation. This could be especially valuable for “turnaround” states, with some areas of strong success. The onus of proof would simply be higher for the additional performance tranche because of country conditions and how they affect project effectiveness. Incentives would improve, both for staff working on difficult countries and for good ministers working under adverse conditions in fragile governments.

The second imperative is to strengthen the general focus on development results, and on the frameworks for measuring and demonstrating them. This requires implementing a short feedback loop, comparable to that now existing in the areas of social, environmental and fiduciary safeguards, to focus attention on the area. While we take no position on the institutional arrangements for doing this, a comparison with the powerful areas of safeguards suggests that both full transparency and independence would be needed to have an effect on incentives. Full transparency might substitute for a mandatory requirement, as it would be difficult for a task manager to publicly justify a decision *not* to seek an external advisory review. Those developing the projects could also benefit in career terms from favorable external reviews, so introducing a welcome incentive into the system.

In the longer run, both of these measures could help to create a stronger and more general link between development results and development financing, including adjusting the PBA formula to place more emphasis on this area.³¹ Linking disburse-

review, to engage with regional quality teams to share them, and to propose a set of experts qualified to contribute to reviews. Participation in an initial advisory review by IEG would provide a useful discipline on later evaluation by preventing second-guessing on the basis of many years of *ex post* information. There are precedents: the Department of Institutional Integrity (INT) both offers advice and conducts investigations.

³¹ In some respects the proposals suggest a partial return to the early approach to country allocations (Mahbub ul Haque “Technical Note of June 21, 1977”). Indicative allocations were based on three

ment to results more broadly would strengthen incentives to develop sound M&E systems, including external validation, as well as enhance the use of full Impact Evaluations where appropriate. It would create incentives within the projects themselves (or the programs supported by the projects) to make this linkage as strong and convincing as possible. Thus, while the proposals take off from the current approaches to shape aid flows to individual countries, they do not replace efforts to develop aid mechanisms that condition flows more directly to outcomes, such as various forms of Output-Based Aid, including COD aid. On the contrary, they provide incentives to build such approaches into country programs wherever possible.

References

Acemoglu, Daron, Simon Johnson and James Robinson 2004. "Institutions as the Fundamental Cause of Long-Run Growth", paper prepared for Philippe Aghion and Steve Durlauf (eds), Handbook of Economic Growth Elsevier, 2005.

Agbonitor Alberto, Anna Pawlowska, Mary Mulusa, Qiang Cui and Soohwha Yi 2009. "IDA Financial Assistance to Fragile States and Performance Trends", Washington DC, December.

Alexander, Nancy 2010. "The Country Policy and Institutional Assessment (CPIA) and Allocation of IDA Resources: Suggestions for Improvements to Benefit African Countries", Heinrich Boll Foundation, Washington, DC.

Alkire, Sabina and Maria Emma Santos 2010. "Acute Multidimensional Poverty: A New Index for Developing Countries", OPHI Working Paper No 38, Oxford Poverty and Human Development Initiative, Oxford UK, July.

Arndt, Channing, S. Jones and F. Tarp 2009. "Aid and Growth: Have We Come Full Circle?", UNU WIDER Discussion Paper 2009/05, October.

Barder, Owen 2010. "Why Do We Give Aid to More than One Country?" Paper presented to Center for Global Development, May.

Birdsall, Nancy and William Savedoff 2010. Cash on Delivery: A New Approach to Foreign Aid. Center for Global Development, Washington DC.

Eifert, Benn and Alan Gelb 2008. "Reforming Aid: Towards More Predictable, Performance-Based Financing for Development", World Development 36(10) October .

European Commission 2005. EU Budget Support: An Innovative Approach to Conditionality, DG Development, Brussels.

equally-weighted components: country performance on growth, country performance on equity, and the Bank's impact at the policy and project levels.

Finkel, Steven, Anibal Perez-Linan, Mitchell Seligson and Dinorah Azpuru 2006. "Effects of US Foreign Assistance on Democracy Building: Results of a Cross-National Quantitative Study", USAID, January 12.

Gelb, Alan, Brian Ngo and Xiao Ye 2004. "Implementing Performance-Based Aid in Africa, Working Paper No 77, Africa Region, World Bank.

Guillaumont, Patrick, S. Guillaumont and L. Wagner 2010. "Augmenting IDA Performance Based Allocation: The Vulnerability Issue". Paper presented to the Annual Bank Conference on Development Economics, Stockholm. FERDI, May 17.

Guillaumont, Patrick 2008. "Adapting Aid Allocation Criteria to Development Goals" FERDI, May.

IDA 2005. "IDA14: IDA's Ratings Disclosure and Fine-Tuning the Governance Factor" International Development Association, Washington DC, September.

IDA 2007. "Selectivity and Performance: IDA's Country Assessment and Development Effectiveness", Paper prepared for the IDA 15 Replenishment, International Development Association, Washington DC.

IDA 2009. "IDA's Performance-Based Allocation and Development Results: An Update", International Development Association, Washington DC.

IDA 2010. "IDA's Performance-Based Allocation System: Review of the Current System and Key issues for IDA 16", International Development Association, Washington DC.

IEG 2009a. Poverty Reduction Support Credits: An Evaluation of World Bank Support, Independent Evaluation Group, World Bank, Washington DC.

IEG 2009b. Annual Report on Development Effectiveness, Independent Evaluation Group, World Bank, Washington DC.

IEG 2009c. The World Bank's Country Policy and Institutional Assessment: An Evaluation, Independent Evaluation Group, World Bank, Washington DC, June.

Kanbur, Ravi 2005. "Reforming the Formula: A Modest Proposal for Introducing Development Outcomes in IDA Allocation Procedures" SAGA Working Paper, January.

Leo, Ben 2010. "Inside the World Bank's Black Box Allocation System: How Well Does IDA Allocate Resources to the Neediest and Most Vulnerable Countries?" Working Paper 216, Center for Global Development, Washington DC, May.

Natsios, Andrew 2010. "The Clash of the Counter-Bureaucracy and Development" Center for Global Development, July.

Ramsey, Cecile 2009. "Fragile and Conflict-Affected States Portfolio Review", Africa Region, World Bank, Washington DC, September

Rajan, Ragurahm, and Arvind Subramanian 2008. "Aid and Growth: What Does the Cross-Country Evidence Really Show?" Review of Economics and Statistics 90 (4).

Wood, Adrian 2008. "Looking Ahead Optimally in Allocating Aid", World Development 36 (7).

World Bank 2003. Breaking the Conflict Trap: Civil War and Development Policy, Washington DC.

World Bank 2004. Strategic Framework for Assistance to Africa: IDA and the Emerging Partnership Model" Africa Region, World Bank, Washington DC.

World Bank 2007 Global Monitoring Report: The Challenge of Gender and Fragile States