At issue: The World Bank policy scorecard: The new conditionality?

Rhetoric says that donors respond to nationally-owned development plans. The reality is that these plans have little impact on policy outcomes or the volume of loans a country receives. Opaque assessments by the World Bank - known as Country Policy Institutional Assessments - do. Critics argue that the scorecards are a way to coerce borrowers into adopting the Bank's preferred model of economic development.

Policy conditions attached to loans simply do not work. Often the prescriptions are wrong for the patient. This can be due to analytical failings or ideological bias in their design. Regardless whether the prescription is right or wrong, a government that does not want to follow the conditions may be able to find a way around them. The debate over the *content* of the conditions continues, but there is consensus that the *mechanism* is broken. Even those conditions which are 'successfully' imposed distort government accountability, undermining the legitimacy of both the policies and the institutions which implement them.

Slower than most, the World Bank and the IMF are realising this. Pitched street battles, lengthy multi-stakeholder reviews and extensive advocacy efforts forced the institutions to change their discourse. Structural adjustment loans became poverty support credits. Ownership became the new touchstone. Civil society was invited to participate in the design of national development strategies. Four years after the introduction of the new nomenclature, civil society organisations said that the fundamental picture had not changed.

In the past year however, a perceptible shift has taken place in the IFIs stance on conditionality. First, an admission on the part of the Fund that it's conditions had extended beyond its mandate and competency. Now grudging agreement from the Bank to review its use of conditionality. Some critics argue that the changes are cosmetic. But if the change does result in the reduced use of conditionality, is it because the Bank and Fund recognise the shortcomings of conditionality? Or is it because conditions make loans unattractive, leading to shrinking portfolios and institutional prestige? Or have the major shareholders simply figured out a better way of getting the policies they want implemented without having to use such ham-fisted methods?

Banking the World Bank way

John needs £50,000 to open a new restaurant. He arrives at his bank, and the loan officer tells him - before looking at his business plan - that his credit limit is only £10,000. But that won't be enough to open the restaurant, he protests. But she won't hear it. The bank has done an assessment of all its customers - not just their financial history, but their education, their employment record, their lifestyle choices. And that is all that John can get. John asks to see the assessment - maybe a mistake has been made. But that, she says, is impossible - "strictly against bank policy". More money might be available, she helpfully suggests, if John were to open a hair salon.

Country Policy and Institutional Assessments

Since the late 70s, at the behest of the deputies of the International Development Association (the Bank's concessional lending window), the Bank has been making assessments of the policies of governments of the poorest countries to guide the allocation of cheap loans. In 1997, came the inception of a new, more formalised system. Since that time, the countries have annually received a country performance rating, known as the 'IDA Country Performance', or ICP.

The ICP is obtained by calculating a weighted average of a policy scorecard (80 per cent) and the Bank's rating of the performance of outstanding loans to the country (20 per cent). This weighted average is then multiplied by a 'governance factor'. The governance factor itself is drawn from the governance-related criteria in the policy scorecard (see box over). The ICP, in conjunction with an assessment of need (based on gross national income per capita) determines the allocation of available funds. What this convoluted calculation means is that the subjective judgement of World Bank economists over the quality of a country's public sector management plays an enormous role in the decision over how much money is available to that country.

The policy scorecard, known as the Country Policy Institutional Assessment (CPIA), is made up of 16 indicators, including those on economic management, structural policies, policies for social inclusion and public sector management. These ratings are prepared annually in all countries by Bank country teams and then subjected to a process of internal review. The exercise takes six months and is estimated to cost $1.5 million. Each criterion is given a score on a scale from one to six.

Lifting the veil

In 2000, the Bank began disclosing the CPIA ratings. But only in an aggregated format that revealed little about the differences between individual countries, why those differences existed, or how the ratings were calculated. The Bank's board members recognised that this secrecy would leave the Bank open to charges that lending decisions were driven by poorly substantiated and subjective decisions, if not outright realpolitik. In 2002, the governors of IDA urged full disclosure of the rating system to "allow it to benefit from open scrutiny". They asked management to report to the board on the readiness of the system for public disclosure at a meeting scheduled for October 2003.

But management dithered on full disclosure. Instead, it proposed that ratings be disclosed in "half-point ranges". Leaked documents from the October 2003 board meeting reveal a split amongst the Bank's directors. While some directors "were concerned that disclosure of IDA's ratings could have a negative effect on foreign investment", others chastised management for taking a "step in the wrong direction" and leaving the "impression that progress was being made when this was not the case." Why the delay?

Revealing in this respect was the discussion at the board around the rating methodology. A "large number of speakers" argued that disclosure should be delayed until the ratings had been improved. The current methodology "was based largely on staff judgement rather than clear, objective measurable indicators". It was decided to put the ratings to an external review process to "assure that the Bank was on solid theoretical footing". This frank admission raises serious questions about the legitimacy of the Bank's past allocation decisions.

An external panel met for two days in Washington in February 2004 to review CPIA ratings and methodology. The panel was made up of nine experts: including 6 from US and European academic or research institutions, the Indonesian deputy minister for international cooperation, and a member of the NEPAD peer review secretariat. The panel made a number of recommendations :

- Simplification of CPIA criteria from 20 to 16;
- Undertake analytic work to better inform the weighting of the various criteria;
- Reconsider the weight given to the 'governance factor', calling this calculation "highly non-transparent" and "excessive in light of the available empirical literature";
- Provide country authorities with an opportunity for comment on the assessments;
- Establish an independent committee to review the CPIA methodology every three years; and
- "Strongly in favour" of full disclosure of the numerical ratings of the 2005 CPIA exercise for IDA borrowers.

The Bank has accepted most of the recommendations of the panel. Importantly, full disclosure of the ratings will start for low-income countries with the 2005 ratings. No progress was made in making the governance factor "simpler and less volatile".

The 'governance factor' contains the following five criteria drawn from the CPIA:

1. Property rights and rules-based governance
2. Quality of budgetary and financial management
3. Efficiency of revenue mobilisation
4. Quality of public administration
5. Transparency, accountability and corruption

Undermining accountability

There are three question marks against the CPIA. The first is over the objectivity of the ratings; the second over their reliability; and the third is whether scorecards should be used at all in aid allocation decisions.

As for objectivity, several of the indicators explicitly reflect an economic bias: the indicator for trade policy, for example, rewards low tariffs, the absence of state marketing boards and the removal of controls on capital inflows. While most of the governance criteria reward behaviour which is lauded across the political spectrum, there is a bias towards such factors as stringent private property rights and low regulation of business.

As regards reliability, a number of indicators, such as those on gender, labour and environmental sustainability, are in areas where the Bank's mandate and expertise is in question. For all the indicators, the creators warn of "substantial margins of error" which mean that cross-country comparisons "should be made with due caution". Small policy changes can have a significant impact on the 'governance factor'. A move to harmonise the scorecards of the different multilateral development banks would further amplify this effect. What if the new harmonised scorecard is wrong?

Do 'good policies' as indicated by the CPIA foster economic growth and poverty reduction? Here there is great controversy and a growing literature. On one side, an "extraordinarily influential" study by Bank economist David Dollar says they do. However, a team of independent economists, lead by a former senior economist at the Bank, William Easterly, was given access to the CPIA database, and concluded that "foreign aid does not raise growth in a good policy environment".

The third question is whether there is a role for *any* kind of scorecard in aid allocation decisions. Opponents of scorecards will rightly argue that they circumscribe policy space and promote external accountability, undermining that of governments to their citizens. Proponents counter that, unlike normal lending, there is a finite limit to the quantity of cheap loans and grants available; in the absence of scorecards, lending will continue to be made, perhaps based on even more opaque criteria. Even if it were possible, giving governments whatever funds they say they need risks punishing the citizens which eventually have to pay for overzealous borrowing. It would also fail to provide the accountability which citizens of donor countries demand of their governments.

Conclusion

The opening up of the CPIA process will provide an opportunity to revisit both its methodological and instrumental validity. Both critics and proponents believe that the 'sunshine effect' of increased transparency will lead to vigorous debate over the accuracy and role of the scorecards.

Before any further steps are taken towards harmonisation, an independent review should be commissioned to examine the relationship between rating criteria and poverty reduction. This review should inform an open debate where borrower country governments and civil society are able to express their opinions on the case both for and against the use of scorecards and what their content should be.

If the use of scorecards survives such a debate, this review suggests:

- *Outcomes-based criteria*, based on a government's ability to improve the lives of its citizens, should take precedence over *policy criteria* which act as a form of ex-ante conditionality. Immediately, it will be pointed out that outcomes-based criteria have been used in the past to justify lending to brutal and/or corrupt regimes. This is why outcomes-based criteria must be accompanied by fiduciary criteria and some criteria measuring respect for internationally agreed human rights.
- A greater role for specialised agencies in the calculation of those ratings which are chosen would lend legitimacy to the exercise.
- Needs assessment done in conjunction with any scorecard might be made more effective if human development measures were combined with the purely income-based indicator used currently. Research into this issue could examine what impact this would have.
- The appropriate weighting of chosen criteria requires further independent, empirical investigation.

Without action on these points, the failed 'one-size fits all' policy prescriptions of the bad old days of structural adjustment lending may simply be reborn in policy scorecards. If this is the case, Bank and Fund moves towards 'reduced conditionality' may prove a boon to lending volumes but a pyrrhic victory for the poor.

Jeff Powell
November 2004

Thanks to those who provided comments on a draft version of this briefing. Any errors are the fault of the author.

For a longer, fully-referenced version of this briefing, see: www.brettonwoodsproject.org/atissuecpia

To comment on this briefing: info@brettonwoodsproject.org

Published by **Bretton Woods Project**
Hamlyn House, Macdonald Road, London N19 5PG, UK
Tel +44 (0)20 **7561 7546**/7
Fax+44 (0)20 **7272 0899**
atissue@brettonwoodsproject.org
www.brettonwoodsproject.org/subscribe
An independent non-governmental organisation supported by a network of UK NGOs and by the C.S. Mott Foundation.