



Conférences
& Séminaires

05

Evaluation and its Discontents: Do We Learn from Experience in Development?

Proceedings of the 9th AFD-EUDN
Conference, 2012

François BOURGUIGNON
Michael A. CLEMENS
James MIRRLEES
Jean-David NAUDET
Leonce NDIKUMANA
Jodi NELSON
Catherine PARADEISE
Ruerd RUBEN
Miguel SZÉKELY

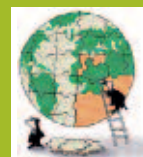


Table of Contents

Acknowledgements	5
Introduction	7
Part 1: Evaluating Development Policies	9
• Chapter 1 : Evaluating Development Policies, by James Mirrlees	11
Part 2: Impact Evaluations: a Tool for Accountability?	23
• Chapter 2 : Impact Evaluations: a Tool for Accountability? Lessons from Experience at AFD, by Jean-David Naudet, Jocelyne Delarue and Tanguy Bernard	25
Part 3: Is Indicator-Based Management a Guarantee of Efficiency?	45
• Chapter 3 : History Revisited : Measurement for Management in Development, by Jodi Nelson	47
• Chapter 4 : How much is Enough? Does Indicator-Based Management Guarantee Effectiveness? by Catherine Paradeise	63
Part 4: Applying Evaluation to Development and Development Aid	93
• Chapter 5 : Dimensioning Development Aid: some Lessons from Evaluation, by Ruerd Ruben	95
• Chapter 6 : Applying Evaluation to Development and Aid: Can Evaluation Bridge the Micro-macro Gaps in Aid Effectiveness? by Leonce Ndikumana	123
• Chapter 7 : Applying Evaluation to Development Policy, by Miguel Székely	151
• Chapter 8 : The Collision of Development Goals and Impact Evaluation, by Michael A. Clemens	169
Authors' Biographies	198

6. Applying Evaluation to Development and Aid: Can Evaluation Bridge the Micro- macro Gaps in Aid Effectiveness?

*Leonce Ndikumana, Department of Economics and Political Economy
Research Institute (PERI), University of Massachusetts^[45]*

Abstract

Donors and governments in aid recipient countries are under pressure to demonstrate the effectiveness of aid, especially given the growing stress on fiscal balances in the context of the global financial and economic crisis. The evidence on aid effectiveness remains mixed at best: while individual targeted aid interventions appear to produce positive results, the impact of aid at the macroeconomic level remains limited. Furthermore, the reporting on concrete outcomes of aid interventions remains inadequate, thus perpetuating doubts as to aid effectiveness. This paper discusses these micro-macro gaps in aid effectiveness and the reporting problem. It proposes some ways in which well-designed and carefully implemented evaluations can help bridge these gaps, and how better reporting and transparency on aid results can advance the aid effectiveness agenda.

[45] The author is grateful for excellent assistance from Theresa Owusu-Danso, doctorate candidate at the University of Massachusetts

6.1. Introduction

There is growing pressure on donors and recipient governments to demonstrate the effectiveness of aid. In donor countries, taxpayers demand tangible proof of the use of the tax money channelled through national aid agencies and multilateral institutions. This pressure has been exacerbated by the adverse impact of the global financial and economic crisis on donors' fiscal balances. At the same time, populations in recipient countries are more and more openly demanding tangible development outcomes, more transparency in the management of aid and better access to reports containing systematic and objective assessments of aid effectiveness. In developing countries, along with increasing democratisation and a free press, governments are facing a growing pool of educated but disfranchised youth demanding genuine improvements in living standards.

These growing pressures and demands for transparency and aid effectiveness are further fuelled by criticisms ranging from analysts arguing that aid has no robust impact on development, to activists who are openly opposed to aid on various grounds. Furthermore, the complexity of the development process makes it difficult to track the impact of aid, which is influenced by many factors emanating from the donor side, the recipient's context and exogenous factors. In addition, the aid industry is a congested market where multiple actors pursue similar goals on the same terrain,

making it even more difficult to sort out the incremental impact of aid interventions.

Nonetheless, evidence shows that development aid has produced substantial positive results at the micro level, whether at project or programme level. Well managed programmes have yielded improvements in school enrolment, access to healthcare, reforms of tax systems and other valuable outcomes. However, at the aggregate level the record remains very mixed, fuelling the debate about overall weak aid effectiveness. Bridging the micro-macro gap remains a critical challenge for the development aid community and national policy makers.

These practical challenges, criticisms, gaps between micro and macro outcomes and domestic political pressures on donors and recipient governments for more transparency on aid call for more effective mechanisms to analyse, monitor, evaluate and disseminate the concrete impacts of aid on development; in other words, there is a call for better aid effectiveness evaluation. While there has been substantial progress in evaluation methods and practice, important gaps remain and there is room for improvement. Moreover, the dissemination of impact evaluation results remains inadequate, which further perpetuates doubts about aid effectiveness. This paper argues that well designed and implemented evaluations, together with better dissemination of evaluation results, can help shed light on these gaps in aid effectiveness. The paper thus emphasises two problems:

the dichotomy between micro-level and macro-level aid effectiveness; and the lack of transparency and inadequate reporting on the concrete impacts of aid. These problems are at the heart of the concerns regarding aid effectiveness in the aid community as well as in recipient countries.

Following this introduction, the paper provides a brief review of the mixed record of aid effectiveness in Section 2. Section 3 highlights the problems at the origin of the micro-macro dichotomy in aid effectiveness. Section 4 discusses the role that evaluation can play in bridging these gaps, and Section 5 concludes.



6.2. Aid effectiveness: a less than stellar record

6.2.1. A backdrop of rising aid volumes

The debate on aid effectiveness and evaluation is taking place in the context of an upswing of aid flows to developing countries. Following a steady decline in the 1990s, total aid by the Development Assistance Community (DAC) member countries has

increased substantially since the turn of the current century. Between 1990 and 2009, total aid to all developing countries by DAC donors rose from USD119.9 billion to 165.3 billion, a 37.8% increase. Based on the trough levels of 2000 (USD84.7 billion), this represents a doubling of the volumes of aid over a decade (Table 4).

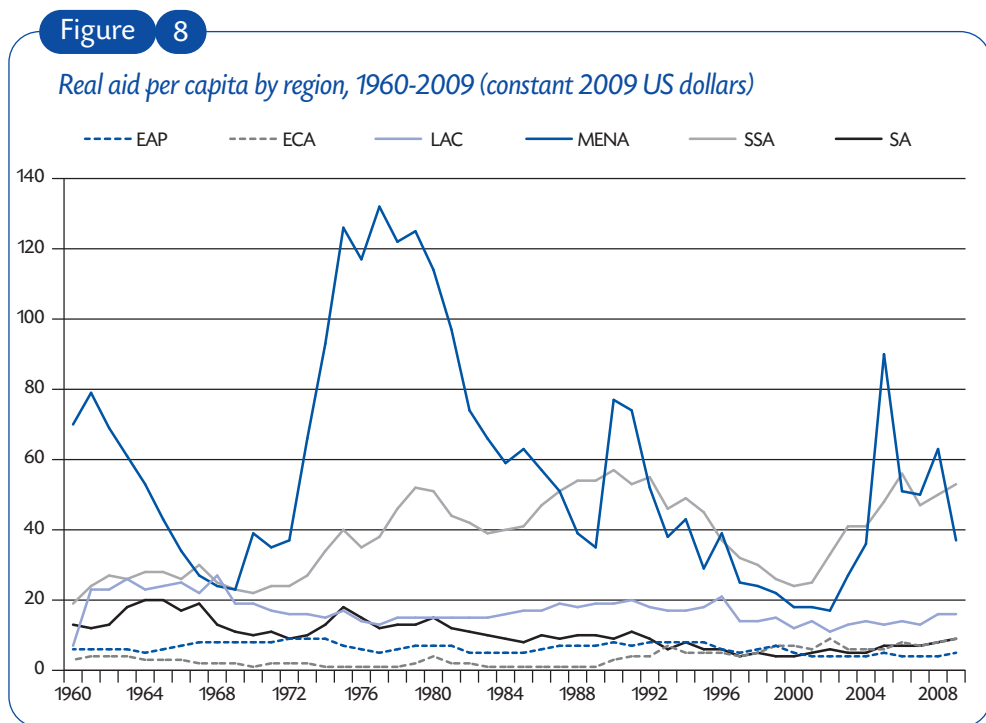
Table 4 *Real annual flows of official aid (constant 2009 US dollars, billions)*

Year	Total	Africa	Latin America	Asia	Europe	Oceania	Unspecified
1960	30.8	9.5	1.6	15.7	2.8	0.2	1.0
1970	44.9	9.3	5.7	18.5	1.0	1.5	1.9
1980	111.4	27.1	5.8	35.3	3.1	2.7	12.8
1990	119.9	41.2	8.6	29.5	2.3	2.3	9.6
2000	84.8	19.2	6.0	19.7	4.6	1.0	11.1
2009	165.3	47.7	9.1	38.6	5.8	1.6	24.9
Change 1990-2009 (%)	37.8	15.8	5.9	30.7	147.6	-26.9	159.1
Change 2000-2009 (%)	95.0	148.7	50.8	95.7	25.0	62.0	122.9

Source: DAC database (online). Nominal values are deflated into real values using the US CPI index.

In per capita terms, the upward trend is most notable in sub-Saharan Africa, Asia and Latin America (Figure 8). The substantial increase in aid to these regions since 2000 has been credited for contributing to high growth in the pre-crisis period.^[46] Aid per capita in sub-

Saharan Africa more than doubled between 2000 and 2009, rising from USD24 to USD54. However, the 2009 levels are still below the peak of USD57 per capita reached in 1990.



Despite the considerable increase in the volumes of aid over the recent years, the quantity of aid remains inadequate relative to the financing needs of developing countries as well as relative to donors’ targets. The High-Level Plenary Meeting on the

Millennium Development Goals (MDGs) held in New York in 2010 under the theme of “keeping the promise” soberly lamented the fact that donors had not kept their promise of increasing aid delivery (UN MDG Task Force, 2011). The report of the MDG Task

[46] Various reports by the multilateral development institutions have listed increasing volumes of aid as one of the key drivers of the high growth in Africa during the years leading to 2008–09. These include the *African Economic Outlook* (by the African Development Bank, OECD, UNECA and UNDP), *UNECA’s Economic Report on Africa*, *UNDESA’s World Economic and Social Prospects*, and the *IMF’s World Economic Outlook*.

Force found that, while official aid had reached a record high of USD129 billion in 2010, this represented only 0.32% of the gross national income (GNI) of DAC members. Only five countries have met the UN target of 0.7% of GNI in official aid.^[47] The report noted a large gap of USD153 billion in 2010 in actual aid delivery. Aid delivery to Africa in 2010 was USD15 billion (in 2004 dollars) below the pledges made in 2005 at Gleneagles (UN MDG Task Force, 2011: 15).

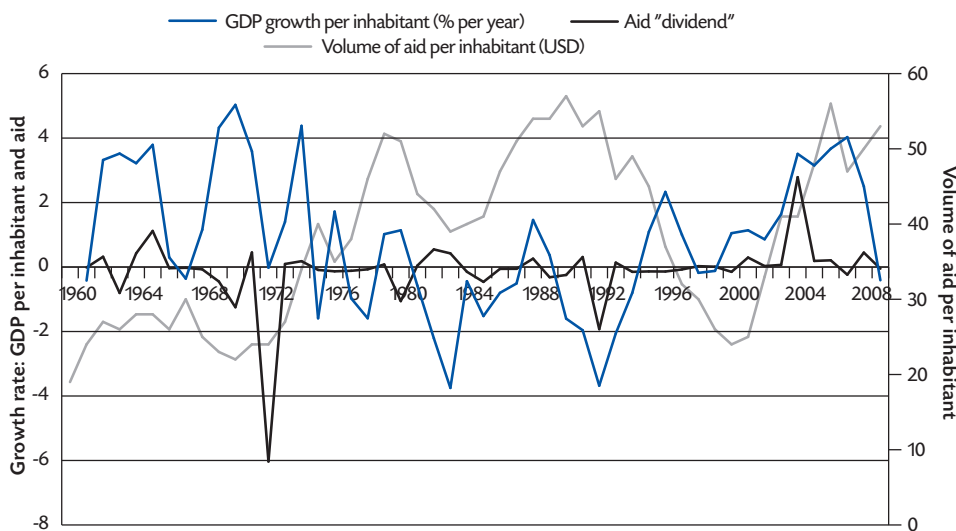
At the same time, developing countries are facing large financing gaps in economic

infrastructure and social sectors. It is estimated that Africa faces an annual gap of USD48 billion in infrastructure financing alone. In 2008, the MDG Africa Steering Group Report (2008) concluded that for African countries to reach the MDGs, public external financing would need to be scaled up by about USD72 billion per year until 2010. Actual disbursements fall far below these targets.

But most importantly, despite the fact that the volumes of aid to developing countries in general have increased over the past years, the record of the impact of aid on

Figure 9

Growth gains from aid in SSA: GDP growth/aid growth



Source: DAC database (online). Nominal values are deflated into real values using the US CPI index. Aid dividend is proxied by the ratio of real GDP growth to the growth rate of real aid per capita.

[47] The five countries are: Norway (1.10% of aid/GNI), Luxemburg (1.05%), Sweden (0.97%), Denmark (0.91%), and Netherlands (0.81%). (Source: OECD-DAC online database).

development remains rather wanting. Growth in sub-Saharan Africa remains below the levels needed to reach national development targets; growth is volatile and has generated inadequate job creation. The credit given to aid for stimulating the recent resurgence of growth in Africa is often exaggerated. Over the long run, the gains from growth are limited. As can be seen on Figure 9, growth elasticity of aid has been low and flat. The recent upswing in aid has only yielded a short-lived spike in “aid dividend”, reverting to a stagnant mean.

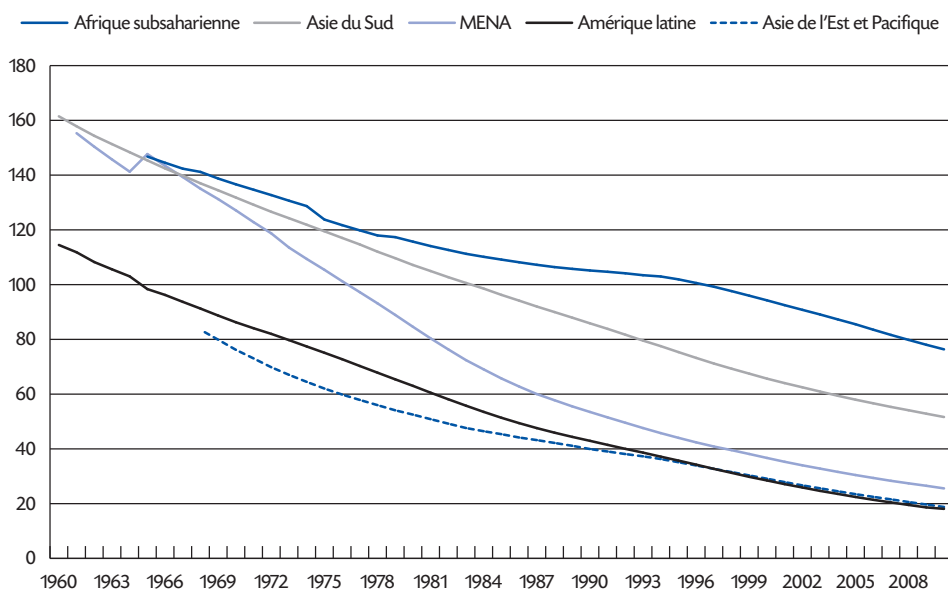
The gains from aid in terms of social development are also less than satisfactory. While aid has supported important projects in education and health, the overall impact

remains inadequate. The deficiencies are most notable in the case of Africa, the region pointed out as having received relatively higher volumes of aid. Infant mortality has declined much more slowly in Africa than in the other developing regions (Figure 10) and many African countries are not likely to reach the MDG target for this development objective.

Although the debate on aid effectiveness has heated up in recent years, efforts to assess the effectiveness of aid date from as far back as the 1960s (Roodman, 2007a). Doucouliagos and Paldam (2005) provide a comprehensive literature review, pointing out the “sad results” of four decades of research on the theme. Over the years, the

Figure 10

Infant mortality by region (per 1,000 live births)



Note: MENA = Middle-East and North Africa.6.2.2. Disputed evidence at the macro level but more encouraging results at the micro level

work on aid effectiveness has been reviewed several times, with conclusions ranging from extreme optimism to quasi-militant pessimism.^[48] But the literature shows a clear demarcation between the findings at the macro level, where the results are mixed, and results at the micro level, where the evidence is much more encouraging regarding the gains from aid. The sections below discuss the evidence at the two levels in turn.

Mixed evidence at the macro level

The ultimate objective of aid is to contribute to improved economic performance and the wellbeing of the population through the provision of financial resources and technical assistance. Naturally, governments and the public from donor and recipient countries expect to see improvements in indicators of economic performance and wellbeing as returns to aid. This explains why the attention in the aid effectiveness analysis has been focused on national level indicators, mainly economic growth, health outcomes and human capital development. Implicitly, the analyst assumes, or rather hopes, that the impacts of aid at the micro level, where the action takes place, somehow translate into macro level impacts at the national level. But the process of aggregation of micro level outcomes into the macro level impacts remains a black box.

The literature on the macro level impact of aid falls into three camps, with a limited number of agnostics along the spectrum of aid effectiveness beliefs: aid works; aid does

not work; aid works under certain conditions (but it works after all). The first camp claims that aid works, and the only concern is that there is not enough aid and that it may not reach the intended recipients. The most vocal advocates in this camp include Jeffery Sachs, who has argued forcefully for a “big-push”-led growth financed by scaling up aid to developing countries (Sachs, 2005).^[49] Sachs and his colleagues argue that external development assistance can help break the poverty trap and that scaling up aid is the exit strategy from the poverty trap. The following excerpt says it all in reference to Africa:

If Africa is caught below the threshold level of infrastructure, and therefore is stuck in chronic low or negative growth, the main policy implication is to raise capital above this threshold ... We propose to increase the capital stock in one step, as it were, through a large, well-targeted infusion of foreign assistance. In other words, we are arguing not for endless flows of increased aid, and not for aid as simple charity, but rather for increased aid as an exit strategy from the poverty trap. For those who fear that aid increases dependency, our response is that aid that is ambitious enough would actually end Africa’s dependency. Moreover, we see no other likely successful strategy for ending Africa’s poverty trap. (Sachs *et al.* 2004: 144)

[48] Key reviews of the literature on aid effectiveness include: Mosley (1980); Hansen and Tarp (2000); Clemens *et al.* (2004); McGillvray *et al.* (2005); Roodman (2007).

[49] See Easterly (2006a) for a critical review of Sach’s argument for a big-push approach to development assistance.

This camp includes analysts who argue that aid has been effective in stimulating growth (see Hansen and Tarp 2000, 2001). Besides the quantitative analysis supporting the aid effectiveness view, there is a large literature from the activist world (NGOs and civil society organisations) calling for a scaling-up of aid.

The second camp argues that aid works, but that it works under certain specific condition. One of the most widely cited studies in this group is by Craig Burnside and David Dollar (2000), which claimed that aid works but only in a good policy environment. This study generated considerable debate and controversies. Some analysts questioned the robustness of the results and the merit of the methodology (see Roodman, 2007*b* for a review). Further investigations refuted the results as too fragile, not robust to sample selection and subject to particular specification of the empirical model (Dalgaard and Hansen, 2001; Easterly *et al.*, 2003). For instance, the interaction of aid and policy and the particular coding of the good policy indicator are found to be the key drivers of Burnside-Dollar results. Outside of academia, the concerns with the Burnside and Dollar proposition were about its policy implications. The proposition implied that aid should go to countries with demonstrated evidence of good policy; that is, aid should be conditioned to good policies. This reopened a can of worms in the debate on aid conditionality. Most fundamentally, it meant that, given that low-income countries and especially those coming out of conflicts also have weak

institutions and policy frameworks, an application of the Burnside-Dollar proposition would leave these countries as aid orphans and trapped into a low-aid-poverty vicious circle.

A number of other analysts have supported the view that aid is effective under certain conditions. Noteworthy studies include Collier and Dollar (2004), who argue that aid effectiveness requires good government institutions. Similarly, Svensson (1999) argues that aid is effective only in democracies. Collier and Dehn (2001) posit that aid can be effective in countries experiencing shocks, but point out that aid effectiveness requires good policies. Patrick Guillaumont and his colleagues argue that aid helps absorb economic and natural shocks, and strongly advocate for allocating official development aid on the basis of economic vulnerability.^[50]

Within this strand of conditional aid effectiveness literature, Dalgaard *et al.* (2004) controversially suggested that aid works outside of the tropics but not within the tropics! The study is empirically fragile, the results being driven by a few countries with specific features, namely Botswana, Egypt, Jordan and Syria (Roodman, 2007*b*). This kind of conclusion feeds the usual deterministic view of development that tends to attribute underdevelopment to fixed factors such as geography. But this view is tenuous; it fails to explain, for example, why geography would prevent Burundi from developing while Switzerland developed, although both countries are landlocked and small.

[50] Guillaumont (2007, 2009, 2010); Guillaumont and Chauvret 2001; Guillaumont and Guillaumont-Jeanneney (2009); Guillaumont and Simonet (2011).

There is a smaller strand of the literature that argues that aid simply does not work, conditionally or absolutely. Rajan and Subramanian (2005) challenge the robustness of the results in studies that conclude that aid works even if it is conditional on good policies. They conclude that aid does not have any statistically consistent effect of growth and that, even in cases where it may exist, the effect is too small to be statistically observable.

William Easterly argues that aid works only if it is well targeted and aligned with individual country's cultural, social and economic conditions (Easterly, 2006*b*). He is critical of large-scale or grand-scheme types of aid interventions, or what he calls a "transformational" approach to aid. The problem is not the money; it is whether the funds are used to meet the specific needs of the intended recipients. Easterly suggests that well-managed aid produces positive results at the micro level in areas such as education and health. He thus favours the "marginal" approach with small-scale targeted interventions (Easterly, 2009).

Overall, the review of the literature suggests that the evidence on the macro-level effectiveness of aid remains mixed with no apparent movement towards any consensus. Now we turn to the micro level impact of aid, where the results are much more promising.

More encouraging evidence on aid effectiveness at the micro level

The key challenge to efforts to document and quantify the effectiveness of aid at the macro level is that macroeconomic outcomes are the result of a multiplicity of

factors, many of which are unrelated to aid, and some of which can affect the effectiveness of aid either positively or negatively. Economies are complex systems where virtually everything depends on and influences everything. Disentangling the impact of a single factor such as aid on macroeconomic outcomes like growth, human capital, health, etc., is a daunting exercise both conceptually and empirically. Moreover, project level interventions may produce macro level results, but these would be observed only several years down the road (Radelet and Banana, 2004).

Moreover, and most fundamentally, aid is only an instrument used to achieve ultimate macro level goals. For the instrument to have an impact on the ultimate goal, a long chain of causalities needs to hold systematically. A chain is as strong as its weakest link; if one node in the chain of causalities fails, then the final result is compromised. For example, the ultimate goal of aid interventions in education is to increase human capital, which in turn would increase growth and generate improvements in the overall wellbeing. So donors finance school construction with the hope that the recipient country will reap future benefits in terms of improved human capital and higher growth. However in practice, for the final result to materialise, not only does aid need to be spent and used diligently, but also agents' behaviour needs to respond appropriately and significantly along the way. So, effectiveness of aid operates at multiple levels and it is the aggregation of the intermediate levels of effectiveness that determines effectiveness at the macro level. Using the example of aid to education through construction of schools, Roodman (2007*a*: 2) summarises

some of the questions that need to be addressed, which points to many ways in which aid effectiveness may be compromised: “Was a school built? Did children come? Did they learn? When they grew up, did they have fewer children of their own? Did they find more rewarding and productive work? Did economic output go up? Did poverty or inequality fall?”

For aid to education to have macro level effects, there are too many “ifs” that need to be satisfied. If more schools are built, school attendance will increase, literacy will increase, households will make more efficient decisions regarding matters relevant for their wellbeing, workers (educated) will be more productive, more output and income will be produced and the living standards will increase. Trying to demonstrate empirically each of these causal statements is a monumental task.

One possible solution to the challenge of demonstrating aid effectiveness is to be less ambitious in the quantitative assessment of aid effectiveness and look not at the macro outcomes but at the micro level outcomes; that is, look at narrower goals. Such an

investigation typically reveals what Clemens *et al.* (2004) refer to as a “micro-macro paradox”: despite the disappointing results at the macro level, there is evidence of successful targeted aid interventions at the micro level. At the sectoral level, aid has also been found to be effective, especially in the areas of education and health. Michaelowa and Weber (2006) find that aid contributes to increasing primary school enrolment. Dreher *et al.* (2007) find similar results. In the area of health, Mishra and Newhouse (2007) find that aid helps reduce infant mortality.

The reality, therefore, is that the aid landscape includes a mixture of successes and failures. The problem is the technology used to aggregate the impact of aid. This is at the root of the fact that the successes have not been able to outshine the failures to produce robust overall positive impacts of aid. There are many reasons for this. Key among these is that the aid enterprise has many structural deficiencies that undermine its effectiveness. These inefficiencies prevent the aggregation of positive micro level results into visible positive outcomes at the macro level.

6.3. Problems leading to the micro-macro paradox

This section discusses succinctly the key structural problems of aid effectiveness that may be at the origin of the micro-macro paradox. The focus is on problems that may be addressed through effective conceptualisation and implementation of evaluation.

6.3.1. A quantity and quality problem

To the extent that aid effectiveness means development effectiveness, then both the quantity and the quality of aid matter. Regarding the quantity of aid, if aid is to produce positive and visible results at the aggregate level, it must reach a minimum threshold. It has been pointed out in several studies and reports that the current levels of aid remain inadequate. They fall short of the investment gaps faced by developing countries and are below the OECD targets of 0.7% of donor countries' gross national income.

Various studies have documented large and even growing financing gaps faced by developing countries. In the case of Africa, for instance, it is estimated that to reach the high growth required to substantially reduce poverty, the continent would need to invest about USD93 billion per annum in infrastructure, including USD41 billion for the power sector (Africa Infrastructure Country Diagnostics, 2009). Currently only USD45 billion are covered, leaving a gap of

USD48 billion, of which USD23 billion is required for the energy sector alone.

For aid to generate meaningful impacts at the macro level, the current levels would have to be substantially increased in a predictable manner to fill the investment financing gaps. Higher and more predictable funding would help achieve higher and less volatile growth, and ultimately foster social development. Improvement in aid effectiveness at the macro level is conditional on increasing the volumes of aid delivery.

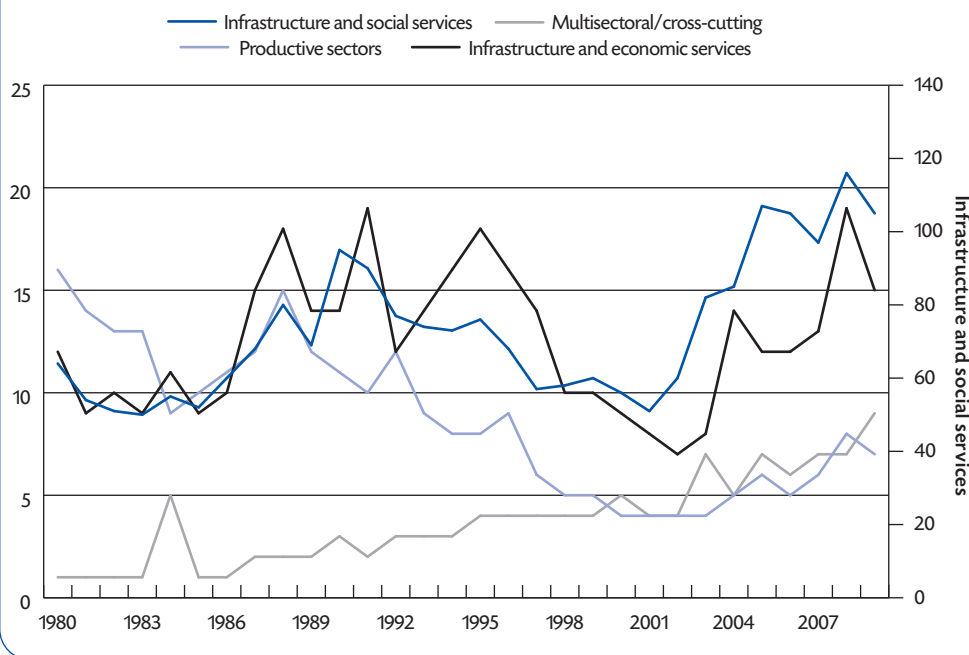
The quality of aid is also essential for aid effectiveness. When aid effectiveness is defined in terms of development effectiveness rather than financial soundness or operational/process conformity, the quality of aid raises a number of issues. The two most prominent issues are allocative effectiveness and predictability. The challenge of allocative effectiveness emanates primarily from the fact that (1) resources are scarce and therefore donors have to make difficult choices on where to invest these resources; (2) there is inadequate information on effective returns to investment in various activities; and (3) there is imperfect evidence on the key drivers of growth at the country level. As a result, aid effectiveness is constrained by the fact that some resources are allocated to sectors with limited returns to investment and with little impact on growth and development.

Allocative efficiency problems are exacerbated by the lack of consistency in the decision-making of donors. Over time, donors change their aid targets and preferences, but it is not always clear whether these shifts are inspired by careful analysis of the expected relative gains from investment in various sectors. So, for example, donors have exhibited a strong bias toward social infrastructure and services, especially since 2000 (Figure 11). In contrast, aid to productive sectors declined since the beginning of the 1990s. As donors focused on poverty reduction as the ultimate goal of aid, the attention shifted to activities and

sectors that were deemed closer to this objective, hence the emphasis on social sectors. This shift, however, is problematic. It has been well documented that sustained poverty reduction requires higher and sustained growth and job creation, which in turn requires adequate investments in productive sectors. Ironically, focussing on the poor by increasing spending on social services has not contributed much to reducing poverty. It is by supporting wealth and job creation through strong, sustained and broad-based growth that sustained poverty reduction can be achieved.

Figure 11

Sectoral allocation of ODA (total all donors, constant 2009 dollars, billions)



Source: DAC database (online). Nominal values are deflated into real values using the US CPI index (obtained from the IMF's International Financial Statistics).

The problem of targeting of aid is amplified by “herd behaviour” among donors, and the tendency to “follow the winner” in a context of high pressure to show results. Individual donors seek to minimise risks by avoiding untapped terrains and focussing on sectors and activities that have gained consensus among the donor community. Moreover, multilateral development institutions, which are key players in the aid landscape, do not have a genuine capacity to set their own targets. They are all accountable to the same governments of member states. As a result, the preferences of dominant donor countries permeate the strategic decisions of multilateral development institutions, so that the preferences of the latter mimic those of the former.

In addition, aid effectiveness is hampered by poor cost-effectiveness, notably due to long and cumbersome aid delivery processes. Aid for seeds and fertilisers is of little help when it is delivered after the end of the planting season. The high cost of aid delivery is also due to ineffective donor coordination in a landscape marked by a proliferation of donors and projects. This increases the burden on recipient governments called to execute, monitor, and evaluate multiple projects and dialogue with multiple donors. It is not surprising, for example, that some recipient governments occasionally call for a moratorium on donor missions during certain periods so they can get down to running their business.^[51]

6.3.2. Weak additionality of aid

The limited record of aid effectiveness at macro level can also be attributed to weak additionality of official development aid. One of the reasons for the weak additionality of aid is that additionality is not integrated in the planning of aid programmes. Additionality of aid can be evaluated at three levels: financial additionality; positive spillover effects into the local economy from aid-funded projects and programmes; and technology and knowledge transfer. Financial additionality of aid stems from the role that aid plays in attracting additional public and private resources. From the public side, aid can crowd in domestic public resources by increasing the capacity to mobilise tax and non-tax revenue. A donor intervention can also crowd in external public resources by enticing other donors to co-fund programmes and projects.

From the private side, aid can play a catalytic role in attracting private financiers or by facilitating public-private partnership funding arrangements. In practice, however, instead of crowding in domestic public resources, aid often tends to have a disincentive effect on tax mobilisation (Ostrom *et al.* 2001: Xviii). Because aid is fungible with other government resources, especially in the case of budget support, high volumes of aid alleviate pressure on the government to mobilise taxes that are politically undesirable.

[51] In Tanzania for example, the government declares a mission-free month during the budget preparation process.

The ability of aid to catalyse additional public and private resources is also limited because this is not explicitly built into aid programming. When institutions make it an objective for their funds to play a catalytic role, aid indeed can crowd in substantial amounts of private finance.^[52] Unfortunately, this practice is not part of the normal business of public sector aid programming among bilateral donors or multilateral development institutions.

Aid projects tend to also have suboptimal spillover effects in the local economy, which limits overall aid effectiveness. Aid programmes often remain virtual islands in the economic landscape, thus minimising their impact at the macro level.

In addition to financing, development assistance can also provide an avenue for the transfer of technological know-how from donors to aid recipients. This in turn would eventually increase productivity in recipient countries, leading to overall higher economic performance. The record of aid effectiveness in this regard is weak. The gains through technology transfer are particularly low in the case of tied aid. Despite calls to move away from tied aid, it still represents a substantial fraction of total aid for many donors, whether *de jure* or *de facto*. This further undermines aid effectiveness.

6.3.3. Failure to influence policy and institutions

There is broad consensus that institutions and good policies are important ingredients

for sustained high long-run growth. Yet, the aid community has not made up its mind whether aid should be used to induce improvements in institutions and policies. For a long time, donors have had it backwards: they have conditioned aid to good institutions and policies. Given that the majority of low-income countries have weak institutions and policies, they then end up receiving less aid. As a result, poor countries are trapped in a stable equilibrium of bad institutions and low growth. Indeed, Birdsall (2007) argues that what is holding African economies in a low-growth high-poverty trap is an “institutional trap”. While growth is believed to be a function of institutions, donors have primarily focused on the direct link between aid and growth, and less on institutions. One of the causes of limited performance of aid in stimulating growth is that little aid has been invested in institution building and that aid has not been sufficiently leveraged to improve the institutional framework in low-income countries.

There are reasons for this limited emphasis on using aid to develop institutions. Some of these constraints are political, whereby donors put their national strategic interests ahead of recipient countries’ economic development goals (Killick, 1998; Kanbur, 2000; Mold, 2009). Thus, bad governance in recipient countries goes unchallenged, and even worse is rewarded by additional aid inflows in the name of national strategic interests. On the recipient side, there is resistance against interventions with an

[52] IFC’s evaluation reports show strong additionality of private sector lending, with the most common form of additionality arising from funds mobilisation. See for example, Independent Evaluation Note No.1, 2008. “IFC’s Experience and Additionality in Middle Income Countries Results and Challenges”; also see IFC (2010) “IFC’s Additionality: A Primer”, September 30, 2010.

institutional emphasis, especially in undemocratic regimes, on the pretence of national sovereignty. Moreover, there is limited knowledge on how exactly to influence the development of good policies and institutions. Donors know good institutions when they see them, but they know less how to engineer them in a particular country. Furthermore, institutions develop very slowly and in a complex fashion. This is not particularly encouraging for typical aid agencies that are bound by short-term “key performance indicators” tied to short-run results. The lack of patience therefore explains the inadequate investment in institution building and in developing the capacity to implement good policies.

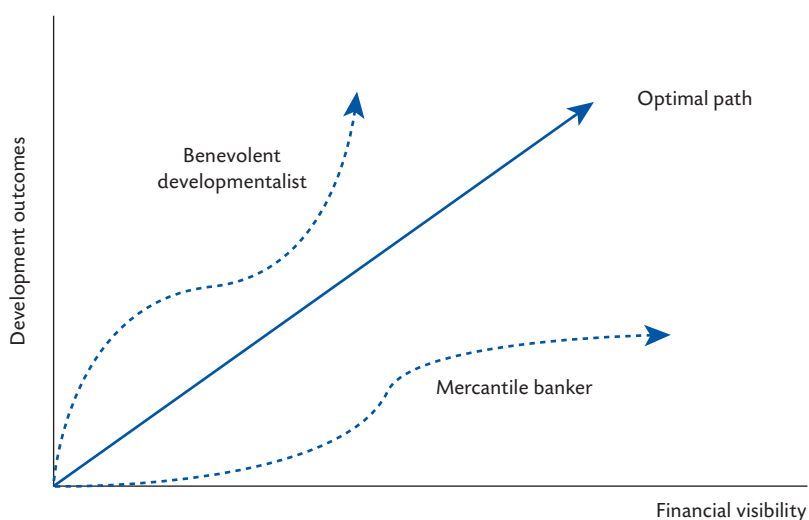
6.3.4. Poor alignment of incentives and interests

Aid effectiveness is also compromised by the lack of convergence between the interests of

donors and those of recipients. Within donor governments and development financing institutions (DFIs), there is often also a lack of consistency between the incentives and interests of the institutions and those of operations officers. Like any investment venture, development aid carries risks. Yet, it is by taking risks that aid can generate the highest rewards in terms of development outcomes. Thus, the donor must strike a balance between financial risk and development outcomes (Figure 12). The tendency of programme and project officers is to err on the safe side, minimising financial risks to demonstrate that money has been used well, thus staying on good terms with internal audit. For development financial institutions, this risk aversion is further motivated by the need to preserve the financial bottom line and good credit ratings. As illustrated in Figure 12, donors face tensions between maximising development

Figure 12

Development-risk trade-off



outcomes, or behaving as “benevolent developmentalists”, and maximising financial viability, or acting as “mercantilist bankers”. In such a context, risk aversion tends to keep aid below the optimal path with regard to development outcomes.

Another way in which incentive incompatibility undermines aid effectiveness is through the pressure to spend the aid budget in the context of the “spend it or lose it” culture of aid budgetisation. This practice induces inefficient allocation of aid resources to activities with little development gains. This causes problems with aid absorption, as large volumes of aid are appropriated to projects and programmes but remain unutilised for long periods. The pressure to “move the money” creates allocative as well as management inefficiencies, especially since programme officers are evaluated not on the basis of effectiveness but on the basis of approvals. Ultimately, these micro level inefficiencies contribute to the overall weak effectiveness of aid at the macro level.

6.3.5. Lack of learning

The challenges described above have been pointed out repeatedly for a long time; yet they continue to permeate the development aid practice. A key reason for this is the imperfection of the learning process in the development aid industry. Evaluation is often not integrated into aid and development policy, and there are inadequate investments in developing evaluation mechanisms. This prevents the development of what Ostrom *et al.* (2001) call the “error-correction capabilities” of systems and institutions that prevent mistakes from generating self-perpetuating inefficiencies. Thus

improvements in aid effectiveness are conditional to the development of effective learning. The question this paper raises is whether evaluation can help fill this learning gap and, as a result, contribute to bridging the gaps between macro level and micro level aid effectiveness. This question is the object of the next section.

6.3.6. Inadequate reporting and dissemination of the concrete impacts of aid

Even in situations where aid has been effective and produced tangible results, especially at the micro and sectoral level, often the results remain unknown to the recipients as well as to the public in the donor countries. It is generally observed that Non-Governmental Organisations do a better job in publicising their interventions and drumming up the results of their projects. This somehow explains the relative success of actions by specialised agencies, and it is consistent with the evidence on aid effectiveness at the micro level discussed above.

The lack of transparency and inadequate dissemination of the results of aid arise primarily from the tradition that aid management is the government’s domain. In developing countries where institutions of public accountability remain under-developed, government operations are not open to the public, and participatory budgeting is not part of the policy and political culture. Thus, while the population is arguably the ultimate beneficiary of aid interventions, it is not systematically informed of the nature of interventions and their concrete results. It is expected that increased democratic consolidation and the

development of a free press will lead to more pressure on governments to open up the aid management process, which will result in better access to information on aid effectiveness for the general public. This is key to building political support for development aid.

Moreover, the technical process of reporting on aid effectiveness remains inadequate and it is not systematically integrated into the

programming and delivery of aid. Even in donor countries where institutions of public accountability are developed, the general public has inadequate access to reports on the results of aid interventions. Reporting to the general public is often defensive, reacting to criticisms from the media and the research community rather than being seen as an inherent obligation of government aid agencies. This tradition undermines the overall aid effectiveness agenda.



6.4. Evaluation can play a major role in bridging the micro-macro gaps

6.4.1. Substantial progress has been made in evaluation practice but challenges still remain

Evaluation is a key component of national aid policy, helping to set goals and access performance. For the USAID (2011: 1), “evaluation is the means through which it can obtain systematic, meaningful feedback about the success and shortcomings of its interventions. Evaluation provides the information and analysis that prevents mistakes from being repeated, and that increases the chance that future investments will yield even more benefits than past investments.” Similarly, the Japanese International Cooperation Agency (JICA) considers that “the primary objective of evaluation is to improve the effectiveness and efficiency of projects by using evaluation results for better planning and implementation” (JICA, 2004).

There has been progress in evaluation methods and practice, especially with the introduction of experimental methods in the design and implementation of evaluation. The use of randomised experiments and randomised control trials (RCTs) holds promise; they are scientific, objective, and minimise sampling bias, thus enhancing the reliability of evaluation results. These methods also have the advantage of being

replicable in various settings (Duflo and Banerjee, 2009; Duflo and Kramer, 2005). The use of control groups enables the analyst to get closer to establishing a causal relationship between a particular intervention and the targeted outcomes.

But there still are many issues, even with the RCT methodology. In particular, the limitations of RCTs stem from the fact that the method works well in situations where an intervention is truly discrete and homogeneous across space and time (Bamberger and White, 2007). This obviously happens in scientific laboratories, but rarely in real social settings. Moreover, RCTs do not completely overcome the perennial problems of attribution of outcomes in a complex system like an economy, where many factors are likely to influence directly and indirectly a particular outcome (Vaessen, 2010). Furthermore, randomisation may face ethical problems as the exercise involves groups that are not benefiting from an intervention that they would otherwise have wished to benefit from. In such a context, it is difficult to explain why some groups would only serve as experimentation objects while others are beneficiaries of the aid intervention under evaluation.

Regardless of the particular methodology used, evaluation faces structural problems.

Key problems are briefly described here. One of the frequent challenges in evaluation is the lack of a clear definition of instruments and targets as well as limited understanding of the exact channels through which the instruments are expected to generate outcomes. This manifests itself in weak logical frameworks of programmes and projects. So, for example, the logical frame of a road project often lists among expected impacts an increase in GDP growth. Certainly, there are many intermediate targets between road construction and GDP growth, and unless the entire chain of causation is clearly defined, it is difficult for evaluation to be effective. Moreover, evaluation faces the classic problem of discrepancy between instruments and goals when too many goals are set with too few instruments. Thus, when the evaluation fails to find the expected outcome of aid, it is difficult to know whether the failure is due to a bad choice of instruments or bottlenecks in the intermediate causal relationships. This also means that innovations such as the so-called “results-based-management” frameworks cannot be effective without a sound definition of targets, instruments and transmission mechanisms.

Moreover, without a good baseline and control groups, reasonable relative progress may be misjudged as failure. For example, in a post-conflict country, large improvements in institutional and economic performance are difficult to achieve in the short run. To illustrate the point, in these countries, achieving the MDGs to reduce poverty by half by 2015 may be impossible. But there may be substantial improvements relative to the no-project scenario. In this case, without

reasonable evaluation criteria, interventions in such settings are inherently set to fail – the “set-to-fail” syndrome. Here the recommendation would be to look at not only the achievement of final targets but also the extent of efforts and relative improvements. To use a sports analogy, evaluation should seek to crown not only most valuable players but also most improved players.

While all donors consider evaluation as an important tool for aid planning and management, they are nevertheless aware of the possible negative repercussions that stem from negative evaluation results. Unsatisfactory evaluations may jeopardise new aid budget allocations (by Congress or Parliament) and even damage the relationships between donor and recipient governments. On the recipient side, there is a risk that negative evaluations may jeopardise new aid. These risks may cause both an underinvestment in evaluation and delayed evaluations due to the “fear-of-the-unknown” effect.

6.4.2. How can evaluation contribute to bridging the micro-macro gap?

Well designed and carefully executed evaluations can help establish better linkages between micro level aid outcomes and the macro level impacts. In other words, to the extent that evaluations are well implemented and systematically integrated along the entire operational cycle, they can help optimise the aggregation of project or programme-level outcomes into national level impacts. This requires a number of innovations in the way evaluations are designed, implemented and

utilised in aid policy. Below are key avenues of possible improvements in that regard.

Evaluation as development diagnostics

Evaluations can help bridge the macro-micro gap to the extent that they are conceived as “development diagnostics” aimed at uncovering the key drivers of intended development outcomes, as well as the channels of transmission from the intervention to the ultimate outcomes. This requires deep knowledge of the sectors involved, the specificities of the country and region, including economic and non-economic features that influence the behaviour of agents and overall economic performance. Evidently such knowledge does not necessarily have to be generated within a particular evaluation exercise. What is needed is close synergies between applied research and evaluation to make evaluation genuinely knowledge-intensive.

Comprehensive ex-ante evaluation as decision-making tool

Many multilateral DFIs have endorsed the practice of *ex-ante* evaluation of development outcomes as a tool for guiding decisions in private sector financing operations. The objective is to identify and attempt to quantify the expected additionality and development outcomes of private sector operations. However, despite the increase in private sector portfolios of DFIs, the dominant lending window remains the public sector. The latter is still not covered by *ex-ante* evaluation of additionality and development outcomes in most DFIs. Two innovations are needed to harness the value added of *ex-ante* evaluations. First, these evaluations need to

be extended to the entire portfolio of multilateral financing institutions including public sector operations. Second, *ex-ante* evaluations need to be more comprehensive and address all aspects of development outcomes, including policy and institutional impacts. At the moment, the analysis on institutional impacts and potential implications for policy is still limited. Yet, this is an area of great potential value added towards improving aid effectiveness and bridging the micro-macro gaps in aid effectiveness.

Evaluation to enhance policy and institutional impact of aid

As discussed earlier, a major weak link in the aid effectiveness chain is the limited contribution of aid to improvements in policy and institutions. Part of the reason is that it is typically not an explicit goal of aid interventions to improve policy and institutions. This is either because institutions and policy are seen as too sensitive or simply because it is believed that aid cannot meaningfully influence policy and institutions. Yet for aid to be more effective at the macro level, or for micro level interventions to translate into national development outcomes, it is indispensable to incorporate improvements of policy and institutions as part of the goals of aid. It is therefore important for evaluations to pay particular attention to the impact of aid on policy and institutions. In this way, evaluations can help the aid process by identifying the factors that make aid effective in improving policy and institutions, and by uncovering the mechanisms and channels that generate such positive impacts. This requires a rethinking of the design and implementation of evaluation frameworks to

bolster the policy and institutional dimensions.

Better integration of evaluation outcomes into operations

The evaluation functions are understandably typically separated from the lending functions of most institutions and governments. This preserves the independence of the evaluator to ensure the credibility and reliability of evaluation findings. However, independence carries some costs. It prevents optimal use of the feedback from evaluation in policy formulation and in the design and implementation of operations. Moreover, the feedback from policy design and operations into evaluations is imperfect. Hence the learning is suboptimal, with the risk that errors and mistakes be repeated over time. An evaluation is good only if it informs policy. One way out is to require that programme officers systematically demonstrate that past evaluation has been integrated in the design and implementation of new interventions. Explicit requirement to build upon lessons from past evaluations would promote the institutionalisation of integration of evaluation into operations.

Transparency, participation and public disclosure of information

For evaluations to serve as an effective tool of aid effectiveness, it is essential to develop a culture of transparency, participation and public disclosure of information in donor agencies and the donor community in general. While most DFIs have established policies on disclosure and access to information,^[53] these policies are rarely known by the target public and they are poorly implemented partly due to lack of adequate resources. Disclosure of information on aid is even less frequent in many government agencies. The increase in the number of donors is accompanied by growing disparities in the practice of information disclosure on aid, despite calls for donor coordination and harmonisation. Yet, public disclosure of information is important to enable the recipient populations as well as the public in donor countries to keep up with the use of aid resources and their concrete impact on development. Thus, transparency and public disclosure of information are key to the aid effectiveness agenda.

[53] The World Bank Policy on Access to Information dates from 2010; The African Development Bank is at an advanced stage of updating its Policy on Disclosure of Information.

6.5. Conclusion

The debate on aid effectiveness has been disproportionately focused on macro level outcomes that look at the impact of aid on national development outcomes such as growth, improvements in the quality of life brought about by better education, and the health status of the general population. However, individual aid interventions do not actually affect these outcomes directly. While aid effectiveness at the aggregate level remains unsatisfactory, the aid landscape contains individual success stories at the micro level. The dilemma is how to bridge this micro-macro gap.

The analysis in this paper suggests that increasing aid effectiveness will require improvements at three levels: (1) improved effectiveness of aid at the micro-level; i.e. at project and sector level; (2) more transparency, better reporting and public disclosure of information on development outcomes; and (3) better aggregation of micro level outcomes into macro level impacts. The paper argues that such an aggregation technology must be knowledge- and institution-intensive. Institutions are critical for not only the quality of outcomes of individual policy interventions (*i.e.* micro level effectiveness), but also for facilitating positive spillover effects of individual interventions into the rest of the economy. Institutions also facilitate learning from experience, which is essential to improvement in policy design, implemen-

tation mechanisms and the overall impact of aid at the macro level.

For evaluations to contribute to bridging the macro-micro gaps in aid effectiveness and help dissipate the clouds surrounding the impact of aid, some innovations in the design and implementation of evaluations, and reporting on aid results are essential. First, this requires a substantive increase in the knowledge intensity of evaluations. Secondly, it is important to achieve higher systematic utilisation of evaluation-generated knowledge in policy and programming than observed in current practice. Third, it is important to improve the reporting mechanisms and systematically embed aid reporting and dissemination of results into the aid planning and delivery processes both at the project and programme levels.

To achieve these innovations, donors and governments in recipient countries need to put their money where their mouth is: if they believe in evaluation, then they must adequately resource it. This requires investing more in evaluation through higher budgetary allocations. It is also necessary to invest more in building capacity and skills in evaluation at the donor and recipient country level. Moreover, it is imperative to develop a culture of transparency, openness and public disclosure of information on aid management in donor and recipient countries. This will improve accountability and ultimately enhance aid effectiveness.

References

AFRICA INFRASTRUCTURE COUNTRY DIAGNOSTIC (2009), *Africa's Infrastructure: A Time for Transformation*, World Bank, Washington D.C, available at www.infrastructureafrica.org

BAMBERGER, M. and H. WHITE (2007), "Using strong evaluation designs in developing countries, Experience and challenges", *Journal of Multidisciplinary Evaluation*, 4(8): 58–72.

BIRDSALL, L. N. (2007), "Do no harm, aid, weak institutions and the missing middle in Africa", *Development Policy Review*, 25(5): 575–598.

BURNSIDE, C. and D. DOLLAR (2000), "Aid, policies, and growth", *American Economic Review*, 90(4): 847–68.

CLEMENS, M., S. RADELET, and R. BHAVNANI (2004), "Counting chickens when they hatch, the short-term effect of aid on growth", Working Paper 44, Center for Global Development, Washington DC. (July).

COLLIER, P. and J. DHEN (2001), "Aid, shocks, and growth", Working Paper 2688, World Bank, Washington DC (October).

COLLIER, P., and D. DOLLAR (2004), "Development effectiveness: What have we learnt?" *The Economic Journal*, 114 (496), F244–71.

COLLIER, P. and A. HOFFLER (2004), "Aid, policy and growth in post-conflict societies", *European Economic Review*, 48(5): 1125–45.

DALGAARD, C.-J. and H. HANSEN (2001), "On aid, growth and good policies", *Journal of Development Studies*, 37(6): 17–41.

DALGAARD, C.-J., H. HANSEN and F. TARP (2004), "On the empirics of foreign aid and growth", *The Economic Journal*, 114(496): F191–216.

DOUCOULIAGOS, H. and M. PALDAM (2005), "The aid effectiveness literature, the sad results of 40 years of research", Working Paper 2005–15, Department of Economics, University of Aarhus, Aarhus, Denmark.

DREHER, A., P. NUNNENKAMP and R. THIELE (2007), "Does aid for education educate children? Evidence from panel data", Mimeo, The Kiel Institute for the World Economy (January).

DUFLO, E. and A. BANERJEE (2009), "The Experimental approach to development economics", *Annual Review of Economics*, 1: 151-178.

DUFLO, E. and M. KREMER (2005), "Use of randomization in the evaluation of development effectiveness", in George Pitman, Osvaldo Feinstein and Gregory Ingram (eds.), *Evaluating Development Effectiveness*, Transaction Publishers, New Brunswick, NJ 2005, pp. 205–232.

EASTERLY, W. (2006a), "The Big Push Déjà Vu, A Review of Jeffrey Sachs's *The End of Poverty, Economic Possibilities for Our Time*", *Journal of Economic Literature*, 44 (March): 96–105.

EASTERLY, W. (2006b), *The White Man's Burden, Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*, Penguin Press, New York.

EASTERLY, W. (2009), "Can the West Save Africa?", *Journal of Economic Literature*, 47(2): 373–447.

EASTERLY, W., R. LEVINE and D. ROODMAN (2004), "Aid, policies, and growth: Comment," *American Economic Review*, 94(3): 774–80.

GUILLAUMONT, P. (2007), « La vulnérabilité économique, défi persistant à la croissance africaine », *Revue Africaine de Développement*, 19(1): 123–162.

GUILLAUMONT, P. (2009), "An economic vulnerability index, its design and use for international development policy", *Oxford Development Studies*, 37(3): 193–228.

GUILLAUMONT, P. (2010), "Assessing the economic vulnerability of small island developing states and the least developed countries", *Journal of Development Studies*, 46(5): 828–854.

GUILLAUMONT, P. and L. CHAUVET (2001), "Aid and performance, a reassessment", *Journal of Development Studies*, 37(6): 66–92.

GUILLAUMONT, P. and S. GUILLAUMONT-JEANNENEY (2009), "Accounting for vulnerability of African countries in performance based aid allocation", African Development Bank Working Paper No.103 (October).

GUILLAUMONT, P. and C. SIMONET (2011), "To what extent are African countries vulnerable to climate change? Lessons of a new indicator of physical vulnerability to climate change", CERDI / Université d'Auvergne.

HANSEN, H. and F. TARP (2000), "Aid effectiveness disputed", *Journal of International Development*, 12(3): 375–98.

HANSEN, H. and F. TARP (2001), "Aid and growth regressions", *Journal of Development Economics*, 64(2): 547–70.

INDEPENDENT EVALUATION GROUP (2008), "IFC's Experience and Additionality in Middle Income Countries Results and Challenges," Note No.1.

IFC (2010), "IFC's Additionality: A Primer", September 30, 2010.

JICA (2004), JICA Guideline for Project Evaluation – Practical Methods for Project Evaluation. (September).

KANBUR, R. (2000), “Aid conditionality and debt in Africa”, In F. Tarp (ed.) *Foreign Aid and Development – Lessons Learnt and Directions for the Future*, Routledge, London, pp. 409–422.

KILLICK, T. (1998), *Aid and the Political Economy of Policy Change*, Routledge/ODI, London.

MCGILLIVRAY, M., M. S. FEENEY, N. HERMES and R. LENSINK (2005), “It works; it doesn’t; it can, but that depends...”, 50 years of controversy over the macroeconomic impact of development aid”, Working Paper 2005/24, World Institute for Development Economics Research, Helsinki (August).

MDGs AFRICA STEERING GROUP (2008), “Achieving MDGs in Africa”, New York, United Nations.

MICHAELOWA, K. and A. WEBER (2006), “Aid effectiveness reconsidered, panel data evidence for the education sector”, Discussion Paper 264, Hamburg Institute of International Economics.

MISHRA, P. and D. NEWHOUSE (2007), “Health aid and infant mortality,” Working Paper 07/100, International Monetary Fund, Washington, DC. (April),

MOLD, A. (2009), *Policy Ownership and Aid Conditionality in Light of the Financial Crisis – A Critical Review*, OECD, Paris.

MOSLEY, P. (1980), “Aid, savings and growth revisited,” *Oxford Bulletin of Economics and Statistics*, 42(2), 79–96.

OSTROM, E., C. GIBSON, S. SHIVAKUMAR and K. ANDERSSON (2001), *Aid, Incentives, and Sustainability, An Institutional Analysis of Development Cooperation*. Sida Studies in Evaluation 02/01.

CLEMENS, M., S. RADELET, and R. BHAVNANI (2004), “Counting Chickens When They Hatch: The Short-Term Effect of Aid on Growth”, CGD Working Paper No. 44, the Center for Global Development,

WASHINGTON DC RAJAN, R. and A. SUBRAMANIAN (2005), “What Undermines Aid’s Impact on Growth?”, IMF Working Paper 05/126, International Monetary Fund, Washington.

ROODMAN, D. (2007a), “Macro Aid Effectiveness Research: A Guide for the Perplexed”, Working Paper 134, Center for Global Development.

ROODMAN, D. (2007b), “The anarchy of numbers: aid, development, and cross-country empirics”, *World Bank Economic Review*, 21(2): 255–77.

SACHS, J. (2005), *The End of Poverty, Economic Possibilities for our Time*, Penguin Press, New York.

SACHS, J., J. MCARTHUR, G. SCHMIDT-TRAUB, K. KRUK, C. BAHADUR, M. FAYE and G. MCCORD. (2004), "Ending Africa's poverty trap", *Brookings Papers on Economic Activity*, 1.

SVENSSON, J. (1999), "Aid, growth and democracy", *Economics and Politics*, 11(3): 275–97.

UNITED NATIONS MDG TASK FORCE (2011), *Millennium Development Goal 8 - The Global Partnership for Development, Time to Deliver*, United Nations, New York.

USAID (2011), Evaluation – Learning from Experience. USAID Evaluation Policy. (January)

VAESSEN, J. (2010), "Challenges in impact evaluation of development interventions, opportunities & limitations", Discussion Paper 2010/01, Institute of Development Policy and Management, University of Antwerp.