

ROOM DOCUMENT 2

ASSESSING THE DEVELOPMENT EFFECTIVENESS OF TOTAL ODA AT THE COUNTRY LEVEL

Item II: iii

This document has been prepared by Anders Berlin, Consultant, on behalf of Sweden, for discussion at the third meeting of the DAC Network on Development Evaluation, 2 – 3 June 2005.

**3rd meeting
2 – 3 June 2005**

Assessing the development effectiveness of total ODA at the country level

At the OECD/DAC Network on Development Evaluation meeting in November 2004 Sweden presented a proposal for a joint pilot evaluation of the impact of the total combined ODA at country level.¹ The proposal was a response to a call for more comprehensive perspectives in evaluation by the DAC Chair, Richard Manning. A working group consisting of Denmark, the Netherlands, and Sweden in collaboration with the DAC Secretariat stood behind it.²

The proposal was favourably received by several participants. The need for a more comprehensive understanding of ODA impact has increased as the architecture of international development co-operation has gradually become more closely integrated. Evaluations of the effectiveness of the combined ODA should be useful to all actors in development co-operation, not least to the developing country governments themselves. Seeking to identify strengths and weaknesses in the development co-operation system, such studies would provide valuable inputs to ongoing efforts to enhance the positive impact of ODA.³

The paper has now been revised along the lines suggested by the November meeting (paper attached to this document). The methodological problems have been underlined as recommended, and the importance of partner country ownership has also been stressed. A brief state of the art note by the economist Arne Bigsten from Gothenburg University has been included as an annex. As pointed out both in the concept paper and in Professor Bigsten's annex there is no established methodology for rigorously assessing ODA impact at country level. However, it is assumed that the proposed study would be worth doing even if every question cannot be answered with a high degree of precision or certainty. A better understanding of the present limits of country level evaluation would be a significant benefit in itself.

The methodological issues will be further examined at a seminar in September 2005. As discussed so far, the seminar has a twofold purpose. First, it will provide the participants with an opportunity to discuss the technical problems of evaluating total ODA with distinguished social science experts. Second, for those who wish to participate in the evaluation more actively, it is an opportunity to join the project task force and contribute to the design of a suitable evaluation process. With the formation of an enlarged task force, the working group will have completed its task. All the members of the DAC Development Evaluation Network will be invited to participate in the seminar and the further evaluation process.

As stressed in the concept paper, the proposed evaluation should be carried out in close partnership between the governments of the countries wishing to participate in the study and their development partners in the DAC Development Evaluation Network. For the enlarged task force the main challenge will be to set in motion an evaluation process that is consistent

¹ Gus Edgren 2005. Assessing ODA at Country Level: Proposal for a Joint Study.

² The task force includes Niels Dabelstein, Evaluation Department, Danida, Denmark, Ted Kliet, Policy and Operations Evaluation Department, Ministry of Foreign Affairs, the Netherlands, Stefan Molund, Department for Evaluation and Internal Audit, Sida, Sweden, and Hans Lundgren of the Secretariat of the OECD/DAC Network for Development Evaluation.

³ The need for comprehensive evaluations of aid effectiveness was recently recognized in the Paris Declaration on Aid Effectiveness.

with established principles of partner country ownership and fully responsive to all its stakeholders.

Interested members of the Development Evaluation Network are welcome to participate in the planning of the September seminar. Comments on the updated version of the evaluation proposal are also welcome.

For comments and further information:

Anders Berlin, Department for Evaluation and Internal Audit, Sida (anders.berlin@sida.se, phone: +46-(0)8-698 53 78)

Stefan Molund, Department for Evaluation and Internal Audit, Sida (stefan.molund@sida.se, phone: +46-(0)8-698 54 43)

Sida/Stockholm
2005-05-12

Assessing ODA Effectiveness at Country Level: Proposal for a Joint Study⁴

Gus Edgren, Stefan Molund, Anders Berlin

1. Background

In his Development Co-operation Report for 2003 the DAC Chairman outlines a fourfold hierarchy of evaluations of aid effectiveness⁵:

1. Evaluation of the impact of the combined support from donors to a particular country. The main question in an evaluation of this kind is how the total aid effort contributes - positively and negatively - to the achievement of desired development outcomes at country level.
2. Evaluation of the effectiveness the development co-operation system itself. In this case the evaluation focuses on questions concerning partnership, the alignment of external support with partner country development strategies and the harmonisation of donor practices.
3. Evaluation of the contribution of an individual donor to the effectiveness of the larger development co-operation effort. The question here is whether and to what extent the individual agency operates in a manner that is supportive of the development co-operation effort as a whole and makes a significant contribution to development results.
4. Evaluation of the internal effectiveness of a particular donor agency. In this case questions concern the institutional effectiveness of an agency rather than its contribution to development impact.

The present paper provides a preliminary thumbnail sketch of what a study focusing on the uppermost levels of this hierarchy might look like. It is noted in the paper that the findings from studies of the effectiveness of the combined ODA may prove useful in evaluations of the programmes and projects of individual donors. But the paper is *not* primarily concerned with such evaluations. It is based on the assumption that a better understanding of the country level effectiveness of total ODA is useful in and of itself.

⁴ This is an updated version of a note by Gus Edgren presented to the DAC Network on Development Co-operation Evaluation in November 2004. It incorporates changes suggested by members of the Evaluation Network. A brief comment on Edgren's original paper by Professor Arne Bigsten, University of Gothenburg has been included as an annex. Bigsten's note includes an overview of recent economic work on the country level effectiveness of development assistance. The present version of the main paper has been discussed in a working group consisting of Niels Dabelstein, Evaluation Department, Danida, Ted Kliet, Policy and Operations Evaluation Department, the Dutch Ministry of Foreign Affairs, Hans Lundgren, the OECD/DAC Secretariat, and Stefan Molund, Department of Evaluation and Internal Audit, Sida. It is intended for further discussion among the members of the Evaluation Network.

⁵ DAC, *2003 Development Co-operation Report*, Section 1.15-16, OECD 2004

The need for a broader, more holistic, understanding of the effectiveness of ODA has increased with recent changes towards a more integrated architecture of international development co-operation. While the objectives of donor agencies are increasingly aligned with the poverty reduction strategies formulated by the governments of developing countries, however, development co-operation evaluations are still largely concerned with project and programme level effects of donor support. This is clearly unsatisfactory. Evaluations that are limited to particular project or programme interventions, or to the contributions of a particular donor, leave out some of the most important aspects of the interaction between the inputs of different actors. Since the effectiveness of ODA increasingly depends on the efforts of a wide range of actors, it has in many cases become impossible, and even irrelevant, to try and attribute specific outcomes to the efforts of individual development actors. In order to establish to what extent ODA has made a difference to the achievement of national development goals it is necessary to review the performance of all the development actors – donor agencies as well as developing country governments and civil society organisations – collectively in a given country context.

Evaluations focusing on the effectiveness of the combined ODA input should be useful to all actors in development co-operation, not least to the developing country governments themselves. Seeking to identify strengths and weaknesses in the overall development co-operation system, studies of this kind would provide valuable inputs to ongoing efforts to enhance the positive impact of ODA.

The present note suggests the first steps towards launching a series of pilot studies in collaboration with interested partner countries. It lays down some principles that should guide the country studies, both regarding the analytical framework and the process of collaboration in carrying them out. Which countries should be studied is a matter to be explored once the framework has been further clarified. The following issues are briefly discussed:

- 1) Assessment of development outcomes
- 2) Assessment of the ODA contribution
- 3) Contributions of individual agencies
- 4) The evaluation process

The experimental character of the study should be stressed from the outset. As explained in the annexed note by Arne Bigsten, there is no established methodology for rigorously assessing ODA impact at country level. Although economists and others have made important advances in recent years many difficult problems remain to be solved.

The proposal is made in full awareness of the complexity of the task. It assumes that the study is well worth doing even if every impact question cannot be answered with a high degree of precision or certainty. Indeed, one of the expected benefits of the study is a better understanding of the extent to which it is presently possible to evaluate the effectiveness of ODA at country level.

2. Assessment of development outcomes

To begin with, the development changes that have occurred in the country over the period of time covered by the evaluation would be identified and assessed. This is an operation in two parts. First the changes are described and, as far as possible, quantified. Second, they are

evaluated from the government's point of view. This is a key point. In line with current views on the importance national ownership and country leadership, the policies and priorities of the partner country government should provide the principal reference points against which change should be identified and assessed.

The assessment should be based on policy statements made by the partner country government to its electorate and local political bodies. Donors often assess change with reference to documents that are in various ways connected with attracting ODA, such as the PRSP or the 'localised' MDGs. Although most governments are indeed committed to the goals laid down in those documents, governments also have to adjust to the constraints of politics. For instance, physical security is almost always a concern that overrides other development targets; in many cases the feasibility of those targets is determined by whether people can feel physically secure or not. Similarly, targets for poverty reduction or plans for redressing imbalances between regions or groups would normally have to be understood as conditional on their peaceful acceptance by the political opposition.

To get a presentation of the national development goals and strategies from the horse's mouth so to speak, it is necessary to engage the government at the political level. If the study covers a period of ten or fifteen years, say, it is important to go back to policy statements and plans that were discussed domestically at the beginning of the period and to recall what were the risks and opportunities that could be seen at that time. Reconstructions done without sensitivity to the political process and to the historical contexts in which goals and strategies were formulated may misjudge the original hopes and intentions of political actors and hence distort the picture of what was really achieved.

Some of the developments can be described with the use of statistical data, such as those regularly presented in economic reports from the government and the IFIs. Such data are regularly reviewed by the partners at political level in Consultative Group meetings and in discussions about the PRSP. In other cases the developments can only be described in qualitative terms. In all cases, the data should be understood as *indicators* of development, and should not be automatically treated as development *outcomes*. To facilitate the evaluation of their significance to the intentions and values of the political actors, it is necessary to relate the quantitative indicators to original plans. If the period of review is long enough, the achievements should include an assessment of increased capacity of national institutions to manage the country's development.

If the first step of the study is undertaken in this fashion, its version of development outcomes may differ somewhat from the most common types of donor reports. (But in countries where these elements are already an important part of the dialogue and the PRS process, the difference may on the other hand be negligible). The presentation of outcomes will place more emphasis on the political context of development than many other reports, and it may also give a better picture of the interaction of economic and political factors in explaining change. If the government feels free to express critical views of its development partners, there may also be criticism of the ways the global system has supported or failed to support the government's efforts. Ideally, this part of the study should be able to stand on its own as the government's account of what has been accomplished during the period under review.

Once development outcomes have been identified and carefully assessed in a government perspective they may also be examined from other points of view. Most evaluations of country performance that have been done so far have been carried out by one agency or

institution, whose value standards have been applied throughout the study. In the present case, we are dealing with a number of donor and partner country agencies, possibly also NGOs, who all have views regarding how identified developments should be interpreted. The different views may converge on broad formulations of objectives, but they may diverge on questions concerning the merits or demerits of specific outcomes, or the trade-off of one development goal against the other. A better knowledge about these differences will raise the quality of development dialogue among the partners.

3. Assessment of the ODA contribution

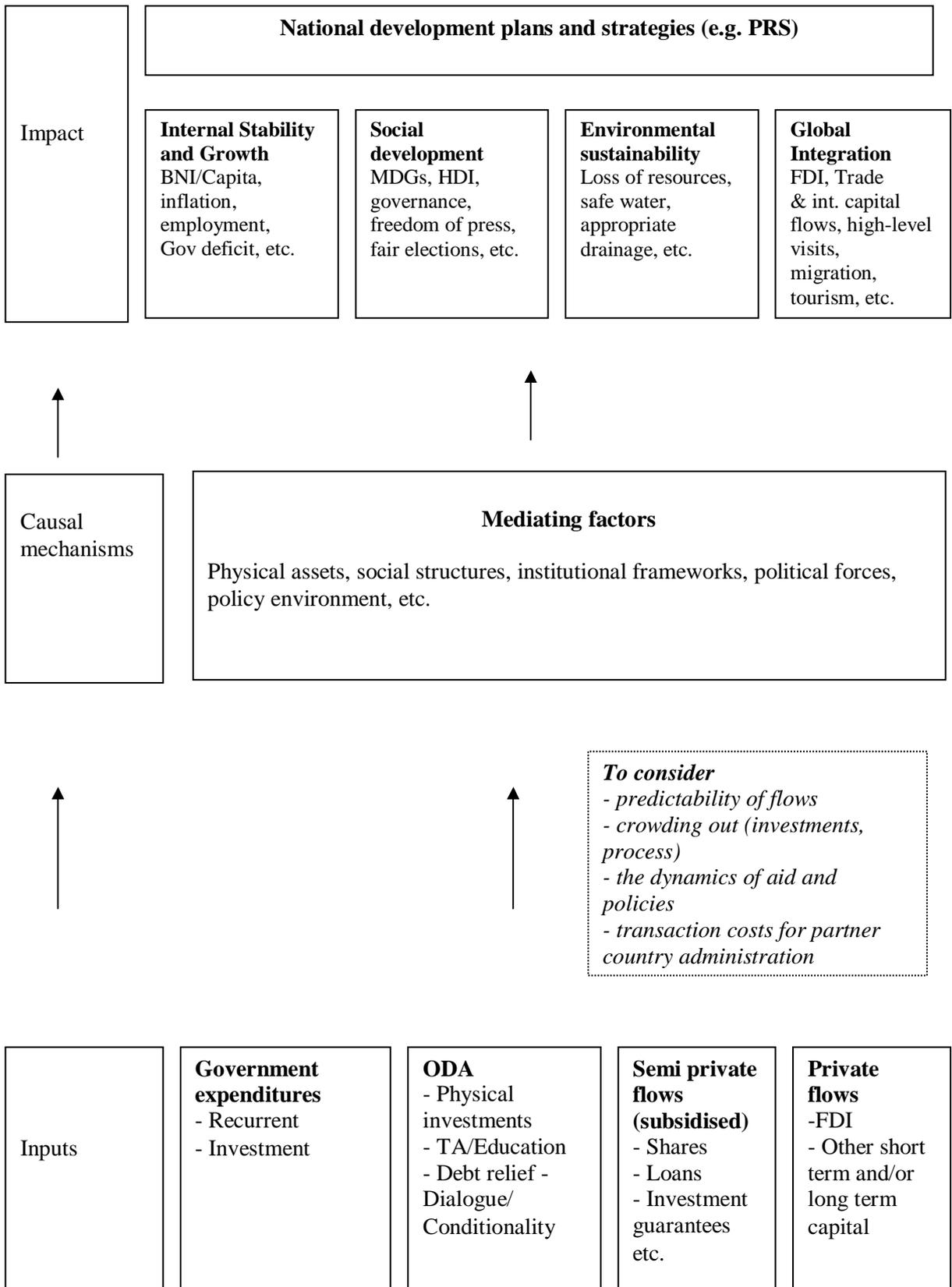
A fairly deep-going assessment of development outcomes is necessary for the discussion of how much and in what way ODA has contributed to developments during the period under review. If one donor has funded infrastructure investment and another has financed school building, this does not tell us much unless we have an idea about the role that these investments have played in enhancing the well-being of people or the economic resources of the nation. The next step in the study will therefore be to assess how effective ODA as a whole has been in contributing to these achievements, together with other sources of funding and innovation, national as well as external. What the evaluation will try to understand is the systemic impact of ODA, its contribution to sustainable development at country level, not just the sum of its project level benefits.

As suggested in the flow chart on page 5 the analysis of the impact of ODA can, with some simplification, be divided into three main components. The first is the mapping of the ODA inputs, the second the assessment of the contribution of those inputs to recorded development outcomes, and the third the identification of causal mechanisms leading from the one to the other. The identification of effects and mechanisms would no doubt be the most difficult of the problems faced by a study of this kind. However, the preliminary step of mapping the ODA inputs is a considerable challenge in itself. Since ODA consists not just of money and hardware, but also of debt relief, technical assistance, and such hard-to-study intangibles as policy dialogue, even to describe it may require its own study. When we proceed to the analysis of impacts and causal mechanisms, questions about the effectiveness of the different ingredients of ODA will arise. Most radically, what is the additional value of delivering ODA in the form of projects and programmes of various kinds instead of just transferring the money?

As tentatively envisaged here, a study of the impact of ODA would begin with an analysis of the allocation of the different ODA inputs over sectors and lines of activity, along with an assessment of the role of the privileged sectors and activities in the achievements of development outcomes. It is quite possible, for instance, that the most important of the recorded development achievements have occurred in the private sector of the economy, whilst ODA has been directed mainly at the public sector and its enterprises.⁶ The overall orientation of ODA investments in terms of sectors, regions and segments of the population is of course highly relevant for the discussion of development strategy.

With regard to ODA effectiveness there are three types of questions, broadly speaking. First, there are questions concerning the value added by ODA, its positive contribution to the identified development achievements. To answer a question of this kind the evaluation must be able estimate with some degree of likelihood what would have happened in the *absence* of the ODA input or with an ODA input of different size or type. What we want to know,

⁶ Edgren, G., *How Effective is ODA in Vietnam?*, UTV Working Paper 2001:1, Sida 2001.



essentially, is whether and to what extent the ODA contribution was a *necessary condition* for the achievement of the outcomes in question, or whether the same outcomes would have been produced anyway. Depending on the case, the answer could be that it was absolutely essential, quite insignificant, or something in between.

For a counterfactual analysis of this kind to make economic sense, it would have to build on clear and specific assumptions about the deployment of alternative funding, in particular domestic resources but in some cases also FDI. A major loss of funds from ODA could spell disaster like in Cambodia and Cuba when Soviet aid disappeared in 1990, but it might also lead to structural reform and reduced aid dependency, as in the cases of Vietnam in 1990 and India in 1970. The counterfactual analysis cannot simply build on fixed input-output relationships but must take into account the political economy at the time.

For example, it would be important to study the relationship between ODA and government revenue. There are cases where it is obvious that easy access to ODA has discouraged domestic resource allocation to a sector⁷, but there are also cases where increasing ODA has led to rising allocations to the sector from domestic resources.⁸ Domestic resource mobilisation in ODA-funded sectors is vital for the financial sustainability of the donor-funded activities and should be carefully examined. Yet, there is no simple relationship between variations in ODA and variations in domestic resource mobilisation independent of the context. Much the same point could be made with regard to the impact of ODA on investments or the current accounts.

A second type of questions concerns the downsides of ODA, its possible negative impact on the ability of the receiving country to achieve its goals. For example, the question of ODA's transaction costs cannot be ignored, although these are difficult to estimate in quantitative terms. Most forms of foreign assistance involve extremely time-consuming and management-intensive procedures of identification, negotiation, review and reporting, which impose a heavy burden on partner country administration.⁹ Donor behaviour that runs counter to the overall objectives of development assistance should be identified, as has been done by the Independent Monitoring Group in Tanzania.¹⁰ The possibility that ODA induces corruption and undermines governance systems would also be considered. .

The third type of questions concerns the determinants of ODA effectiveness rather than the effects of ODA. More concretely the questions in this case concern the factors that in a particular country context may enhance or, alternatively, stand in the way of greater ODA effectiveness. Among the contextual factors explaining the relative effectiveness of ODA some are external to the development co-operation narrowly conceived, while others are part of it. For example, in developing countries where cost-efficient production in agriculture or minerals is feasible opening up of markets in the developed countries would often have more important consequences for development finance than the ups and downs of ODA. More

⁷ See Dollar, D. and Pritchett, L. about the 'Zairean Disease' in *Assessing Aid: What Works, What Doesn't, and Why*, OUP New York 1998.

⁸ Agbonyitor, A.D.K, *Development Expenditures and the Local Financing Constraint*, World Bank, 1998, <http://econ.worldbank.org/docs/787.pdf>

⁹ Brown et al, *Aid Transaction Costs in Vietnam*, ODI, London, December 2000; Hauge, A., *The Role of Monitoring and Evaluation Capacity Development in Improving Government Performance: Lessons Learned from the Uganda Poverty Reduction Support Process*, WB-OED, August 2003.

¹⁰ ESRF, *Enhancing Aid Relationships in Tanzania: Report of the Independent Monitoring Group*, Report submitted to the Consultative Group for Tanzania, Dar es Salaam December 2002

generally, it would be important to consider the coherence of all donor country policies towards their developing country partners.

4. Contributions of individual agencies

The macro approach to studying the effects of ODA could give valuable indications of where ODA has been effective in a strategic sense, and it would also give an idea of the financial sustainability of ODA-supported developments. It will provide insights that could be useful for public investment planning and for the PRS process. If one is to descend in the evaluation hierarchy suggested by the DAC Chair to the level of the contribution to development impact of individual donor agencies (the ‘third level’ above), the selection of projects and programmes for deeper analysis will have to be guided by the findings of the macro study rather than by individual agency interests. Unless they can be related to the country level analysis, individual projects will be of little interest.

One way to satisfy the interest of individual donors in third level analysis would be to try and find linkages between donors’ own effectiveness studies and the macro analysis suggested above. For instance, if the macro-level analysis suggests that too little emphasis has been given to regional disparities, or that donors have used too many external consultants when local expertise is available, the same questions could be raised in effectiveness studies of individual donors.

The macro-level analysis would also offer an opportunity for in-depth studies of certain sectors where donors have been much engaged. While it is usually difficult to attribute specific development *outcomes* to a given resource input, the *process* of achieving these outcomes could be subject to an interesting evaluation of individual donor contributions. Evaluations of this kind could be undertaken as separate exercises following the macro-level analysis, and they could be undertaken by individual agencies or by several agencies jointly.

However, as noted above, the purpose of the present initiative is not primarily to facilitate donors’ efforts to trace the effects of their own contributions. In the proposed study one could well consider the adequacy of different aid modalities and modes of co-operation promoted by particular donors or groups of donors. However, this would be done in order to learn lessons of a more general kind regarding what works and what does not work, rather than to assist individual donors report on the results of their particular contributions.

5. The evaluation process

As pointed out by way of introduction, the proposed study is an experimental one. It is not quite clear to what extent it would be possible to answer the overriding question concerning the effectiveness of ODA in supporting partner country development strategies. Even if the answers provided by the study would turn out to be less precise or less certain than expected, however, the effort would be worth while. A better understanding of the difficulties of evaluating the effectiveness of ODA at country level would be useful in itself.

Thus, it is suggested that a pilot study be undertaken to test the technical feasibility of the concept outlined in this paper. The points below on the evaluation process try to make the evaluation practicable by commissioning it to independent expertise, while at the same time keeping it open for all partners to take part.

Partnership. The evaluation should be undertaken in close partnership with the governments of the selected case study countries. In previous discussions among the members of the

OECD/DAC Network on Development Co-operation Evaluation the participants agreed that a study of this kind could not and should not be initiated without strong partner country support. The implications of this for the governance and organisation of the evaluation would have to be further discussed in a wider task force. However, one would expect the participating partner countries to be closely involved at all stages of the evaluation.

Country coverage. It would seem practical to restrict the evaluation to a small number of countries, say two or three. As the role of ODA varies greatly from one country to another, however, one would also want the sample to reflect some of that variety. For example, it would probably be useful to include a country that is highly dependent on ODA along with one where ODA is less important. However, the question of country selection is a multifaceted one that cannot be fully resolved until the study framework has been further elaborated.

Donor coverage. It would help if the donors actively involved in the study together represented a sizeable share of ODA to the partner countries under study, but it is not absolutely necessary. It is more important that the study is regarded as useful by a large majority of the concerned donors, and that all or most of them are able to support it with information as required.

Period of review. In order for development outcomes to be clearly identifiable, both in the field of economic resources and in capacity development and political institutions, it is preferable to choose a relatively long period like ten or fifteen years. During such a long period, aid flows may have varied significantly, which may call for using sub-periods in the second-level analysis.

Terms of reference. Detailed terms of reference along the general lines suggested in this paper should be agreed by all the partners engaged in the study. The focus should be on assessing the country's development results in relation to national priorities and on establishing what role to be attributed to the external partners in achieving them. The study should suggest and discuss alternative strategies towards the same goals, for donors as well as for the Government.

Analytical perspectives. The choice of analytical approach is likely to be a consequential one. It is important that the analytical models adopted are fully adequate to the scope of the study and the entire range of questions that it raises. Presumably, a combination of perspectives from economics, political science, social anthropology, and other disciplines would be required. Yet, trying to decide what would be the optimal mix of expertise for a study of this type would clearly be premature at this early stage.

Building on existing studies. An evaluation like the one envisaged here must make full use of existing research work, including evaluations produced by donor agencies; this goes without saying. Although evaluations of the combined effectiveness of total ODA are practically non-existent, there are evaluations of the country-level programmes of individual donors that might provide useful inputs to the study. The desk synthesis of existing research work will provide valuable guidance for any subsequent fieldwork. For the evaluation community a synthesis of existing evaluations would be valuable in itself.

Commissioning of the study. The study should be commissioned to a group of independent experts, chosen by the partner country and the external donors together. The research team

should include suitable mix expertise with members from the partner country as well as from the donor countries. It could also included researchers from institutions that are independent of either side.

Stakeholder consultations during the research phase. Since the evaluation is intended to advance the scope of the development partnership, the research process should be undertaken with the partners commissioning the study involved. Interim reports should be discussed with national and international stakeholders in workshops and seminars. The aim of the stakeholder discussions would not be to produce a common value standard, not even a shared view of what has been accomplished, but rather to exchange views on the reported outcomes and to help the research team get a deeper understanding of the development processes.

Joint assessment of results: When the expert group has completed the final report, it should be assessed by the sponsoring partners together. If requested by some of the partners, the first discussion could be held *in camera*, but the partners should agree at the outset on the eventual purpose to discuss the findings in some form at a more inclusive forum of Government and donor agencies as well as NGOs. The usefulness of further deepening or widening the study might also be discussed in this forum.

Annex 1

Gothenburg
050410

Assessing Aid Impacts

Arne Bigsten

There is at present a discussion among donors about the possibilities of doing comprehensive aid impact evaluations at the country level. At the request of Sida Edgren (2004) has written a report on how one could evaluate the impact of the combined support from donors to a particular country, that is how one would go about analysing how total aid contributes to the achievement of desired development outcomes at the country level. This note provides some further thoughts on this issue.

Edgren's proposal

Edgren argues that in order to establish to what extent ODA has contributed to the achievement of national development goals it is necessary to review the performance all development actors – donors and partner country agencies – collectively in a given country context. He further argues that one needs an analysis of the overall development cooperation system. He suggests a few pilot studies according to the following principles:

(1) *Assessment of development outcomes.* The evaluation should be made against the priorities of the government. The outcomes are also to be evaluated against the standards of existing partnerships.

(2) *Assessment of the combined ODA contribution.* The evaluation shall try to understand the systematic impact of ODA, its contribution to sustainable development at the country level, not just the sum of its project level benefits. The analysis shall begin with a review of the allocation of resources from ODA between different sectors and lines of activity, along with an assessment of the role that the privileged sectors and activities have played in the achievements. It is emphasized that also indirect effects need to be taken into account. Edgren suggests that there are three questions to consider: First, what is the value added by ODA, its positive contribution to recorded achievements. This requires that formulation of a counterfactual. The formulation of such a counterfactual is the key difficulty in the whole analysis. It is further suggested that one should take political economy aspects into account, look at the relationship between ODA and government revenue and domestic resource mobilisation, as well as non-material resources such as management. Second, one should look at the negative effects of aid, for example in the form of increased transaction costs. Third, one should look at what factors enhance aid effectiveness.

(3) *Contributions of individual agencies.* The selection of projects and programmes with a certain donor should be guided by the results of the macro study.

(4) *Proposal for a pilot study*

Country coverage – 2-3 countries with different degrees of aid dependence.

Donor coverage – a number of them but not necessarily all.

Period – 10-15 years.

TOR – “The focus should be on assessing the country’s development results in relation to national priorities and on establishing what role to be attributed to the external partners in achieving them.” This is the main task of the study, and it is a hard one to do in a convincing way.

Commissioning of the study – independent researchers.

Building on existing studies - although evaluations of the combined effectiveness of total ODA are rare, some studies of country-level programmes of individual donors have been analysed.

Stakeholder consultations during the research phase. It is argued that the research should be done in close collaboration with as many agencies as possible. There should be interim reports to be discussed with stakeholders in workshops and seminars to help the research team.

Joint assessment of results – the end product should be assessed by the sponsoring parties together.

To conclude, Edgren’s proposal is to do a study of the overall impact of total aid on individual countries. The key methodological question is to what extent one can identify the separate the effect of aid from the effects of other influences.

The Aid Effectiveness Literature

The aid evaluation business is multifaceted with evaluations ranging from purely qualitative reasoning to hardboiled econometric work. I will here start with a brief review of recent work.

Project evaluation

There exist established techniques to analyse the returns to individual stand-alone projects. Cost-benefit analyses have been done for the last forty years, and a mass of project evidence has been collected. These have shown, for example, that aid-financed projects in Africa have often had high returns as noted by the Commission for Africa. The most rigorous project evaluations are done by the World Bank, and for the period 1993-2002 they report an average return of 22%, which is very high (CfA, 2005, p. 298). It is also noted that the projects have been characterised by improving sustainability and better institutional impact.

When it comes to evaluating “process project” aiming to develop institutions, it seems clear that they are so far removed from the final development outcomes that it is hard to do a quantitative analysis of impacts on final targets such as growth or poverty reduction..

The results of project evaluations have generally seemed reassuring, but in the 1980s there was a concern about the micro-macro-paradox (Mosley, 1987). The question raised was why aid projects seemed to work, while it was hard to find positive aggregate growth effects of aid.¹¹ The dissatisfaction with the quality of macro analyses of aid impacts lead in the 1990s to the emergence of a literature analysing the aid-growth link at the macro-level.

Cross-country regressions

¹¹ See also White (1992) and Cassen et al (1994) for further discussion.

The literature from the late 1990s has used cross country regressions to identify conditional relationships between aid and growth. A key paper is Burnside and Dollar (2000), who argued that aid is more effective in good policy countries, while its growth impact was negligible in bad policy environments. Their policy recommendation was that aid should be concentrated to good environments. Later Collier and Dollar (2002) argued that aid should be concentrated in high-poverty countries if one wants to have the maximum impact on global poverty reduction. There are difficult econometric issues involved in the estimation of these effects, and Hansen and Tarp (2001) criticised some of the choices made by Burnside and Dollar and found in their own analysis that aid can have a positive impact on growth also in less favourable policy environments.¹² Easterly, Levine, and Roodman (2004) find that the results in Burnside and Dollar are not robust.

There is now a certain degree of agreement that aid has had a positive growth impact, but Hansen and Tarp (2001) showed that there are decreasing returns to aid in terms of its growth impact. They explain this by constraints on absorptive capacity, which mean that the positive effect of aid on growth peters out once its size exceeds a certain proportion of GDP. Others have pointed at Dutch Disease effect, that hinders export growth. McGillivray (2005) finds in his survey of the literature that various studies place the diminishing returns in the range 15-45% of GDP. One possible policy implication of this may be that the use of donor money is to try to remove bottlenecks can have high returns.

There is an abundance of studies showing that good policy is good for growth. Still, donors have an exaggerated view of their ability to change the recipient's policy, if the recipient government is not itself willing to do so. Burnside and Dollar (2000) found that aid does not systematically influence policy. Past donor behaviour has created a severe time consistency problem. Alesina and Dollar (2000) found that in general policy is quite persistent and hard to change by aid. They also found that donors respond quickly to democratic change, but not to economic policy change. Devarajan et al (2000) found that sustained policy changes cannot be brought about unless the government is committed, but that finance increases the benefits of reforms when a country actually undertake them. Aid increases confidence in the reforms and leads to higher investment, and it also makes it possible for the government to finance public services that are complementary to private investment. Aid also increases the likelihood that reforms will be sustained. Dollar and Svensson (2000) looked at 220 economic reform programmes and found that the outcome of reforms can be predicted quite well by economic characteristics existing before the reforms. Knack (2000) found that there was an adverse effect of aid, which during the period 1982-95 led to increased corruption and a deterioration of bureaucratic quality and the rule of law. Aid thus seems to be a weak policy instrument to reduce corruption. So while aid has little impact on governance, governance has a large impact on the effectiveness of aid.

There are many studies that show that macroeconomic shocks have negative effects on growth and poverty. Terms of trade shocks are very important in Africa. Booms are not translated into sustained income increases, while negative shocks tend to lead to sustained losses. Aid can increase growth in shock prone countries, particularly if donors are willing to "insure" recipient countries against the shocks. Collier and Dehn (2001) showed that donors are not

¹² Reviews of the debate are provided by Collier and Dollar (2004) and Dalgaard, Hansen, and Tarp (2004). Recently, Clemens, Radelet, and Bhavnani (2004) addressed the issue by defining 'short impact aid' to distinguish aid for which a growth response may be expected. A recent paper by Dalgaard and Tarp (2005) reports high returns to foreign aid in a cross-country analysis.

willing to compensate countries for the effects of price shocks, while they are willing to compensate for natural disasters.

A range of other aspects have also been discussed. Lensink and Morrissey (2000) showed that aid has a smaller positive effect on growth when it is more volatile. Collier and Hoeffler (2004) show that aid can enhance both growth and security post-conflict. Both aid and policy affect the economic variables that determine the risk of civil war.

Country studies

Some comprehensive country studies have been done. There were for example a set of studies done by SASDA in the early 1990s (for example, Bigsten et al 1994a, 1994, on Tanzania and Zambia) and another set of country studies done by White (1994) on macroeconomic impact of aid. Some country studies have looked at the impact of structural adjustment, that is considered policy impacts in addition to the amount of aid and how it is spent. White (1999) report on a series of country studies done to evaluate the impact of programme aid funds. Here they try to trace how programme aid has affected macroeconomic aggregates such as imports and government spending, and from this they try to infer an impact on growth. It is acknowledged that it is very hard to establish an appropriate counterfactual at the country level. White (1999, p. 93) writes for example, with regard to the analysis of the impact of funds and policy, that “econometric estimation of meaningful economic relationships at the required level of disaggregation is virtually impossible”. What is applied is an “ad hoc approach which start by identifying aid in a consistent set of internal (government finance) and external accounts (balance of payments). From that starting point a ‘no aid’ counterfactual set of accounts are constructed.” From this set of data one then attempts to get an idea about aid impacts. These impacts refer to factors that are taken to be important for growth, but one cannot identify the direct impact on growth or poverty reduction. These types of studies all suffer from the problem of defining a counterfactual in a reasonable way.

Summing up

Gunning (2004) distinguishes four types of studies in the aid effectiveness literature. “The first one, which drew both on theoretical and empirical work, argued that conditionality as practiced by most donors (an offer of aid in exchange for promises of specific policy changes) did not work (see World Bank, *Assessing Aid*, 1998). The second one, closely related, stressed that in many (but not all) uses of aid fungibility undermined that donors intentions: aid increased money for other uses than the project ostensibly financed by aid. The third one used growth regressions to investigate determinants of aid effectiveness such as the quality of the recipient’s policy. ... The idea is that the effectiveness of aid is conditional on a wide variety of country characteristics. These three strands came together in the World Bank (1998) *Assessing Aid* report. The first two strands demolished the paternalistic case for aid, the idea that the donor could influence government policies and project design and management in the best interest of the recipient country. The third strand, as already noted, is still surrounded by econometric controversy, but the conclusion that aid works better in good policy environments is widely accepted.” The fourth strand of the literature discusses how aid should be allocated across countries to achieve the best poverty reduction impact.

McGillivray (2005, p. 2) concludes from his survey of the aggregate country-level impact of aid that it does have a positive growth impact, but that there is still room for criticism with regard to fungibility, insufficient alignment between donors and recipient government policies, commercial tying, proliferation of donor activities within recipient countries, and insufficient policy coherence.

The main effort in the literature during the last ten years has been to try to measure the impact of aid on incomes by using aggregate panel data. The processes that mediated the outcomes were left in a black-box. The result of this literature is that aid in general has a positive effect on growth and thereby poverty reduction, at least in good policy environments. There is still a heated debate about what a good policy environment is, though. The studies have also given some information about the relative importance of various growth determinants. But this line of research tells us little about why and how aid works. So there is a need to implement impact analyses that looks at different modalities of aid. The donor community needs to know what interventions work.

How to Do Country Studies

The composition of aid has changed in recent years towards more of budget support and less traditional project support. To evaluate the impact of this type of general aid one needs to evaluate the whole recipient system. There is also a new drive to dramatically expand aid flows to Africa (in particular), and if that comes about it seems very important to understand on a more disaggregated level what kind of aid works and why.

The outline by Edgren suggests that one should undertake impact analyses at the country level. Impact analysis can mean several things. One is that one undertakes a rigorous analysis using a counterfactual, while others may interpret it to mean studies with a focus on outcomes, or analysis done several years after the project is finished, or country or sector-wide studies generally. To get high quality results one needs a counterfactual, and new techniques have emerged to identify credible control groups such as impact randomization and propensity score matching (Ravallion, 2001). The problem with these techniques is that they are only appropriate for certain types of interventions, namely those that can be seen as delivering treatment to a clearly defined group. They are not appropriate to policy reforms or programme aid, and not to technical assistance or institutional development. So if we are looking for comprehensive country impacts they do not solve the problem. The reason for the popularity of the cross-country panel data approach is that it makes it possible to control for a range of variables. But it is a problem that it assumes that all countries have the same underlying structure. And apart from that there is the problem that the approach does not tell us anything about what happens in the black box.

The objective of aid could be defined in terms of economic growth or poverty reduction or achievement of the MDGs. If this is the aim of a certain country, we need to understand what the constraint on the achievement of those aims are. Discussing the impact of aid we need to show how aid, e.g. in the form of financial support to the government, helps countries to overcome these problems.

So it seems fair to say that with regard to aid evaluation there is a gap between the macro work, such as regressions to assess aid effectiveness, and micro work analysis. There is a debate concerning whether comprehensive studies of individual countries can be sensibly done. So what types of studies can be done? Let us look at some alternatives.

Computable general equilibrium models

These kinds of models have the advantage that they have an inbuilt counterfactual and take indirect effects into account. However, they have some limitations if one wants to answer the question about what aid inflows does for growth. Since the models typically optimize within constraints, aid inflows lift constraints and thereby lead to higher output. This is virtually

guaranteed. But the models are stylized versions of the world that don't take all aspects into account, and it may not be a credible proof that "aid works" if one can show within such a model that it leads to higher output. The models may be more useful to test different types of policies and then compare outcomes.

Econometric time-series country studies

The reason that aid-growth literature is very much dominated by cross-country or panel studies is that those studies are much easier to do than a time series study of one country. To do that one needs good time series data and then run unit root tests, try different specifications and estimation techniques. The risk of getting unusable or insignificant results is very high. The problem with this type of country specific studies is due to degrees of freedom constraints and difficulties in finding cointegrating relationships. Still, it could be worth to test whether such analyses can produce usable results.

Mixed Methods Country Studies

White (2005) provides two arguments for country case studies.¹³ The first is that although they abandon representativeness they allow for more for depth. A case study can go deeper into some issues that cannot be addressed effectively in a desk-based portfolio review. The second argument is that for example the approach used by the OED of the World Bank is not based on case studies alone but combines a number of approaches. It is thus a mixed approach, and this is probably what Edgren has in mind.¹⁴

¹³ White (2005) argues that there is scope for meta studies, which aggregate performance from different studies. He notes the such analysis involves six steps, namely: (1) define the problem; (2) collect the relevant studies to be reviewed; (3) screen the studies for quality; (4) standardize the indicators and code them; (5) produce a synthesis measure; and (6) present results.

¹⁴ The World Bank regularly does Country Assistance Evaluations of Bank assistance programmes to specific countries, but these do not evaluate the recipient's overall development process. These studies start by defining a set of objectives, typically a sub-set of the Client's development objectives. It is noted that "an unsatisfactory assistance program outcome does not always mean that Bank performance was also unsatisfactory, and vice-versa. This becomes clearer once we consider that the Bank's contribution to the outcome of its assistance program is only part of the story. The assistance program's outcome is determined by the *joint* impact of four agents: (a) the Client; (b) the Bank; (c) partners and other stakeholders; and (d) exogenous forces (e.g., events of nature, international economic shocks, etc.). Under the right circumstances, a negative contribution from any one agent might overwhelm the positive contributions from the other three, and lead to an unsatisfactory outcome."

Bank evaluations have three dimensions (triangulation) in their evaluation to check for consistency in the results. They are "(a) a *Products and Services Dimension*, involving a "bottom-up" analysis of major program inputs-loans, AAA, and aid coordination; (b) a *Development Impact Dimension*, involving a "top-down" analysis of the principal program objectives for relevance, efficacy, outcome, sustainability, and institutional impact; and (c) an *Attribution Dimension*, in which the evaluator assigns responsibility for the program outcome to the four categories of actors."

"In evaluating the outcome (expected development impact) of an assistance program, OED gauges the extent to which major strategic objectives were relevant and achieved, without any shortcomings. Programs typically express their goals in terms of higher-order objectives, such as poverty reduction. The country assistance strategy (CAS) may also establish intermediate goals, such as improved targeting of social services or promotion of integrated rural development, and specify how they are expected to contribute toward achieving the higher-order objective. OED's task is then to validate whether the intermediate objectives produced satisfactory net benefits, and whether the results chain specified in the CAS was valid. Where causal linkages were not fully specified in the CAS, it is the evaluator's task to reconstruct this causal chain from the available evidence, and assess relevance, efficacy, and outcome with reference to the intermediate and higher-order objectives.

Evaluators also assess the degree of Client ownership of international development priorities, such as the Millennium Development Goals, and Bank corporate advocacy priorities, such as safeguards. Ideally, any differences on dealing with these issues would be identified and resolved by the CAS, enabling the evaluator to

A crucial question is what the transmissions mechanisms are from aid to growth. They could conceivably be through investment, import, public sector fiscal aggregates, and government policy. One could look at a set of aid-financed projects and analyse how aid goes through the recipient budget. McGillivray and Morrissey (2001) refer to a number of fiscal response studies of individual countries. It is desirable to try to establish impact that open the black box to help us understand what works and what doesn't. One could possibly start from what we know from cross-country studies, and make a checklist of issues to look at for at the country level. If a comprehensive country study is to be done one would need to consider both policy impacts and aid project impacts.

Some types of aid interventions are very hard to evaluate in quantitative terms. For example, it seems reasonable to argue that aid should be used to remove bottlenecks leading to aid absorption problems and inefficient administration. The analysis of the effects of such attempts on economic growth or poverty reduction are essentially lacking, largely because it is very hard to obtain credible results. Chauvet and Collier's (2004) have looked at the impact of aid on policy and institutional reform, but the analysis is so far rather tentative. Burnside and Dollar (2004) find strong evidence that institutional quality determines the effectiveness of aid. There is no attempt to look at the implications of the reforms for growth and poverty reduction.

Limitations of Country Studies

One could also consider whether direct aid is more effective than other forms of interventions. The recent commission reports (Sachs and Blair) do discuss these issues, but otherwise the underlying analysis has been rather weak. The problems that aid is supposed to solve are often implicit.

The question about the relative usefulness of other forms of interventions is problematic. One would, for example, need to take a view on whether the best way to improve access to the markets in developed countries is to give export supporting aid or reduced import restrictions in the North.

Externalities related to international public goods may be important motivations for aid. This can be agricultural research, aids research or the maintenance of peace. The literature does not tell us much about the role of aid in the provision of global public goods. And it does not measure the negative external effect of a civil war on neighbouring countries. All this are probably very important, although it is hard to put a value on the effects, and it is hard to incorporate such analysis in a country study..

An important characteristic of African economies is that economic life is risky. This drives the behaviour of economic agents and make them choose safe but possibly not the most productive activity combinations. There is not much of a reinsurance market in for example

focus on whether the trade-offs adopted were appropriate. However, in other instances, the strategy may be found to have glossed over certain conflicts, or avoided addressing key Client development constraints. In either case, the consequences could include a diminution of program relevance, a loss of Client ownership, and/or unwelcome side-effects, such as safeguard violations, all of which must be taken into account in judging program outcome.”

OED utilizes six rating categories for **outcome**, ranging from highly satisfactory to highly unsatisfactory. It also considers various dimensions of institutional development impact and sustainability.

African agriculture, which suggests that there is a market failure which the donors at least temporarily could help remedy by sharing the reinsurance risk. This might well be a worthwhile type of intervention, but not much has been done as yet that can be analysed.

There is also macro risk due to terms of trade swings and the like, and Guillaumont and Chauvet (2001) find that aid is particularly efficient in economies vulnerable to trade shocks. It has been argued also that growth may be retarded by shocks when countries are unable to borrow. Therefore there have been insurance programmes, such as STABEX, which have compensated LDCs for export price shocks. Measures of this nature do have a role to play. So this may be a case for insurance rather than direct aid transfers.

There is also the large question about what donor countries should do with regard to their overall policy. The coherence of different policies was discussed in the Swedish Globkom-report, where it was concluded that we should consider the impact of all our policies on world poverty. Trade policy may be more important than aid policy. This is hard to evaluate empirically, and it would probably not be part of a comprehensive country evaluation of aid impacts.

References:

- Alesina, A., Dollar, D. (2000), "Who Gives Aid to Whom and Why?", *Journal of Economic Growth* 5(1): 33-63.
- Bigsten, A. et al (2994a), *Evaluation of Swedish Development Co-Operation with Tanzania: A Report for the Secretariat for Analysis of Swedish Development Assistance*, SASDA, Ds 1994:113.
- Bigsten, A. et al (1994b), *Evaluation of Swedish Development Co-Operation with Zambia: A Report for the Secretariat for Analysis of Swedish Development Assistance*, SASDA, Ds 1994:114.
- Burnside, C., Dollar, D. (2000), "Aid, Policies, and Growth", *American Economic Review* 90:847-68.
- Burnside, C., Dollar, D. (2004), "Aid, Policies and Growth: Revisiting the Evidence", World Bank Policy research Working Paper No 3251, Washington DC.
- Cassen, R. and associates (1994), *Does Aid Work?* Oxford University Press, New York.
- Chauvet, L., Collier, P. (2004), "Development Effectiveness in Fragile Sates: Spillovers and Turnarounds", Centre for the Study of African Economies, Oxford University.
- Collier, P., Dehn, J. (2001), "Aid, Policy, and Shocks", Policy Research Working Paper 2688, World Bank, Washington, DC.
- Collier, P., Dollar, D. (2002), "Aid Allocation and Poverty Reduction", *European Economic Review* 46(8): 1475-1500.
- Collier, P., Guillamont, P., Guillamont, J., Gunning, J.W. (1999), "Reforming Stabex", *The World Economy* 22:669-82.
- Collier, P., Hoeffler, A. (2004), "Aid, Policy and Growth in Post-conflict Societies", *European Economic Review* 48(5): 1125-45
- Commission for Africa (2005), *Our Common Future*, March.
- Clemens, M.A., Radelet, S. Bhavnani, R. (2004), "Counting Chickens when they Hatch: The short-term of aid on growth", Working Paper no. 44, Center for Global Development, Washington DC:.
- Dalgaard, C-J., Tarp, F (2005), "The Return to Foreign Aid", Discussion Paper 05-04, Institute of Economics, Copenhagen University.
- Devarajan, S., Dollar, D., Holmgren, T. (eds.) (2000), *Aid and Reform in Africa*, The World Bank, Washington DC.
- Dollar, D., Svensson, J. (2000), "What Explains the Success or Failure of Structural Adjustment Programmes?", *Economic Journal* 110: 894-917.

- Easterly, W., Levine, R., Roodman, D. (2004), "New Data, New Doubts: A comment on Burnside and Dollar's 'Aid, Policies and growth (2000)', *American Economic Review* 94(3): 781-84.
- Economic commission for Africa (2005), *Our Common Future*, March.
- Edgren, G. (2004), *Assessing the Contribution of ODA to National Development Results: Proposal for a Joint Study*, mimeo, Stockholm.
- Guillamont, P., Chauvet, L. (2001), "Aid and Performance: a Reassessment", *Journal of Development Studies* 37:66-92.
- Gunning, J. (2004), "Why Give Aid?", Paper for the 2nd AFD-EUDN Conference: Development Aid. Why and How, Paris, November.
- Hansen, H., Tarp, F. (2001), "Aid and growth regressions" *Journal of Development Economics* 64(2): 547-70
- Knack, S. (2000), "Aid Dependence and the Quality of Governance: Cross-Country Empirical Tests", mimeo, Development Research Group, World Bank, Washington, DC.
- Lensink, R., Morrissey, O. (2000), "Aid Instability as a measure of uncertainty and the Positive Impact of Aid on growth", *Journal of Development Studies* 36(3):31-49.
- McGillivray, M. (2005), "Is Aid Effective?", mimeo, WIDER, Helsinki.
- McGillivray, M., Feeny, S., Hermes, N., Lensink, R. (forthcoming), "It Works; It Doesn't; It Can But that Depends ...;50 Years of Controversy over the Macroeconomic Impact of Development Aid", *Journal of International Development*.
- Millennium Project (2005), Investing in Development. A Practical Plan to Achieve the Millennium Development Goals. Report to the UN Secretary General, UN, New York
- Mosley, P. (1987), *Overseas Development Aid: Its Defence and Reform*, Wheatsheaf, Brighton.
- Ravallion, M. (2001), "The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation", *World Bank Economic Review* 15(1): 15-40.
- White, H. (1992), "The Macroeconomic Impact of Development Aid: A Critical Survey", *Journal of Development Studies* 28(2): 163-240.
- White, H. (1994), *The Macroeconomics of Aid. Case studies of four countries*, SASDA, Ds 1994:115
- White, H. (1999), *Dollars, Dialogue, and Development: An Evaluation of Swedish Programme Aid*, Sida Evaluation Report, Sida, Stockholm.

White, H. (2005), "Challenges in Evaluating Development Effectiveness", in Pitman, G.K., Feinstein, O.N., Ingram, G.K. (2005), *Evaluating Development Effectiveness*, World Bank Series in Evaluation and Development Vol. 7, Washington DC.

World Bank (1998), *Assessing Aid. What Works, What Doesn't, and Why*, Oxford University Press, Oxford.