

Unclassified

COM/DCD/TAD(2010)2

Organisation de Coopération et de Développement Économiques  
Organisation for Economic Co-operation and Development

11-May-2010

English - Or. English

DEVELOPMENT CO-OPERATION DIRECTORATE  
TRADE AND AGRICULTURE DIRECTORATE

COM/DCD/TAD(2010)2  
Unclassified

## HOW TO EVALUATE AID FOR TRADE: APPROACHES, METHODOLOGIES AND PROCESSES

7 June 2010, OECD Conference Centre, Paris

*This paper is submitted for COMMENT to the joint meeting of the DAC and Working Party of the Trade Committee for COMMENTS under item 3 of the agenda [COM/DCD/TAD/A(2010)4/PROV]*

*This work is all about determining what the aid-for-trade community needs to know to improve the evaluation of aid for trade. This paper outlines the results chain and how different impacts should be considered at the outset in the planning of projects. It discusses the application of impact evaluation techniques to aid-for-trade activities. While selecting appropriate tools is important, a number of processes in evaluation such as independence, the use of joint evaluation and involvement of partner countries are essential to ensure meaningful results.*

Contact:

William Hynes - Tel: +33(0)1 45 24 15 68 - Email: [william.hynes@oecd.org](mailto:william.hynes@oecd.org)

Jean-Jacques Hallaert - Tel: +33(0)1 45 24 13 82 - E-mail: [jean-jacques.hallaert@oecd.org](mailto:jean-jacques.hallaert@oecd.org)

JT03283436

Document complet disponible sur OLIS dans son format d'origine  
Complete document available on OLIS in its original format

English - Or. English

**TABLE OF CONTENTS**

EXECUTIVE SUMMARY ..... 3

HOW TO EVALUATE AID FOR TRADE: APPROACHES, METHODOLOGIES AND PROCESSES .. 4

    1. Introduction ..... 4

    2. How are we proceeding? ..... 5

    3. Approaches and Methods ..... 6

    4. How should the process of evaluation change?..... 11

    5. How do we know what works at the global level?..... 12

    6. Applicability..... 14

    7. Conclusions and Next Steps ..... 14

**Figures:**

FIGURE 1: MCC IMPROVING ROAD QUALITY LOGIC MODEL ..... 7

FIGURE 2: INTERVENTION LOGIC FOR ROADS ..... 8

FIGURE 3: EVALUATION PROBLEM ..... 9

## EXECUTIVE SUMMARY

This work is all about determining what the aid-for-trade community needs to know to improve the evaluation of aid for trade. How should the aid for trade community engage with their evaluation counterparts? Earlier discussions have shown that there are no universal approaches to evaluating the broad and diverse range of activities which are part of the Aid-for-Trade Initiative. Better use of existing generic and development guidance on evaluation seems to be most needed. Against that background it is recommended to provide clear evaluative guidance to policy and management staff in the formulation and implementation of aid-for-trade programmes.

The scoping note [COM/DCD/TAD(2009)3] focused on preliminary issues such as definitions, levels of evaluation, donor activities and approaches in evaluation and the conceptual challenges they face. This paper outlines the results chain and how different impacts should be considered at the outset in the planning of projects. It discusses the application of impact evaluation techniques to aid for trade activities. While selecting appropriate tools is important, a number of processes in evaluation such as independence, the use of joint evaluation and involvement of partner countries are essential to ensure meaningful results.

Before we can say anything about the impacts of aid for trade on trade and economic growth, we need to get the micro-interventions right, we need to gather the right data, adapt evaluation practices and work together with other donors and partner countries. The paper suggests ways to use evaluative information at the micro level to tell a broader story at the macro level through meta-evaluations, synthetic evaluations and econometric analysis.

The final paper will bring the scoping note and this paper together and complement these reflections with a comprehensive overview on how to evaluate aid for trade. The paper will incorporate the findings of the Meta-evaluation on Aid for Trade being undertaken. This will provide insights for future work on country case stories and studies as well as providing key insights into what aid-for-trade related evaluations currently tell us and how evaluation practice could be amended to improve the evaluation of aid for trade.

This work forms part of the *performance assessment framework* output result of the joint DAC-TC 2009-2010 PWB [COM/DCD/TAD(2008)7] that looks at how to better evaluate aid-for-trade projects and programmes. This work complements the work on Management for Results in aid for trade. These two work streams and the meta-evaluation will converge into a single final output, a guidance document on how to improve the performance assessment of aid for trade.

## HOW TO EVALUATE AID FOR TRADE: APPROACHES, METHODOLOGIES AND PROCESSES

### 1. Introduction

1. There is a growing need for the Aid-for-Trade Initiative to demonstrate that projects and programmes were well implemented and achieved their development objectives. The Initiative has helped to mobilise more aid for trade but this has only increased the pressure to show not only that the Initiative is a success in terms of the process it has established, but more importantly it has led to positive changes in partner countries in achieving their trade and development goals. To do this we need to know a lot more about what works and what doesn't.

2. At a time when aid budgets are under renewed pressure and scrutiny, more rigorous evaluations are needed. This is especially true of aid for trade which has become an increasingly important priority in development cooperation and has attracted increasing aid flows. While this is a positive development, it comes with it increasing pressure to demonstrate that these resources have been well spent. But the promise and possibilities of evaluation should be kept in perspective. Evaluation concepts and tools have proliferated in recent years creating in some minds a "fixation with measurement", an "audit culture", "bulging toolboxes loaded with frameworks and concepts".<sup>1</sup>

3. This work is all about determining what the aid-for-trade community needs to know to improve the evaluation of aid for trade. How should the aid for trade community engage with their evaluation counterparts?<sup>2</sup> Earlier discussions have shown that there are no universal approaches to evaluating the broad and diverse range of activities which are part of the Aid-for-Trade Initiative. Better use of existing generic and development guidance on evaluation seems to be most needed. Against that background it is recommended to provide clear evaluative guidance to policy and management staff in the formulation and implementation of aid-for-trade programmes.

4. This note provides an update on ongoing OECD work on the evaluation of aid for trade. It builds on the scoping note [COM/DCD/TAD(2009)3] presented at the November joint meeting of the DAC/WPTC on aid for trade and the presentations from bilateral and multilateral agencies on their evaluation efforts. This note explores the results chain some donors have developed for transport infrastructure and asks whether impact assessments are applicable in this area. It also notes a number of

---

<sup>1</sup> As discussed at the future of development series at Sussex University. More details can be found here: <http://www.oxfamblogs.org/fp2p/?cat=27>

<sup>2</sup> The DAC Evaluation Network is preparing a guide on Development Evaluation Resources and Systems including profiles on all Network Members. It contains an analysis of current trends and challenges in evaluating development assistance. Core elements of the evaluation function are covered, including: the role of evaluation in the development agency, funding and human resources for evaluation, independence, quality assurance, coordination, mechanisms to ensure use of evaluation findings, the role of partner countries and support to partner country evaluation capacity development.

processes which would enable better evaluation, such as independent, joint evaluations conducted with the participation of partner country representatives.

5. The rest of the note proceeds as follows;

- Section 2 provides an update on recent developments in aid-for-trade evaluation.
- Section 3 discusses approaches and methods for evaluation.
- Section 4 examines how changes in the process of evaluation practice could improve the learning function and value of evaluation.
- Section 5 outlines ways in which evaluation findings at the micro level (project level) could be used to get an impression of the macro picture (national/regional trade and development outcomes).
- This is nuanced by Section 6 which discusses applicability.
- Section 7 concludes

## **2. How are we proceeding?**

6. Evaluation has been a recurring theme in OECD's aid-for-trade work programme. This section reviews progress to date and describes next steps. After a first scoping note in 2007 and discussion during the 2008 OECD Policy Dialogue on Aid for Trade, the current work focuses on how to evaluate aid-for-trade programmes and projects. This work is undertaken in close cooperation with members, representatives of international organisations, and evaluation experts within OECD. The scoping paper outlined the background for evaluating aid for trade, reviewed past evaluation of trade-related assistance, established a set of objectives and principles to guide work on evaluation, and asked a number of questions on the approaches which might help take the work forward.

7. The discussion at the last joint DAC/WPTC meeting was preceded by a number of presentations from agencies that have been actively engaged in the evaluation of trade related assistance.<sup>3</sup> The aim was to (i) share experiences on conceptual issues; (ii) explain evaluation methodologies and tools; and (iii) highlight lessons learned from past efforts to evaluate trade related activities. The discussion showed that donors use varying definitions, objectives, approaches and resources. Members thus concluded that common evaluation frameworks were not required for aid for trade. At the meeting of the Evaluation Network in February 2010, it was clear that the evaluation community also felt that no new guidelines were needed for aid for trade; generic approaches were considered sufficient. The evaluation community has seen a proliferation of specially tailored guidelines for different areas with little added value. On the other hand, a discussion concerning the conceptual challenges, in particular the linkages in the results chains, the attribution problem, the sequencing of different programmes and projects could bring greater operational clarity.

8. Aid for trade is not new and evaluations are undertaken as part of regular donor activities. However, evaluations in areas which have a potentially important trade dimension do not always investigate or consider these potential trade impacts. This is partly because of the low profile of trade as an

---

<sup>3</sup> the presentation is available here:

[http://www.oecd.org/document/59/0,3343,en\\_2649\\_34665\\_44077883\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/59/0,3343,en_2649_34665_44077883_1_1_1_1,00.html)

objective of development cooperation but also because of the perceived methodological challenges. Yet there is an urgent need to generate new knowledge about aid for trade interventions. Evaluation, ex post assessments of effectiveness and impact, confirming project expectations and measuring impacts are essential to generating evidence and insights on what works.

9. The OECD Aid for Trade Community Space provides an inventory of planned evaluations which was developed by the DAC Evaluation Network Secretariat based on donor's forward planning. The list is provisional and indicative only. Members are encouraged to discuss the planned evaluations that are relevant to them with their evaluation colleagues. This will stimulate further interaction and improve communication between the aid for trade and evaluation communities.<sup>4</sup> Such interaction is essential because many issues that the aid-for-trade community is grappling with are the ones shared by the wider development and evaluation community.

### 3. Approaches and Methods

10. The weak trade performance of low income countries is in part due to poor trade and transport infrastructure. A lack of transport infrastructure is a key binding constraint that aid for trade aims to address. Djankov, Freund, and Pham (2008) show that trade delays reduce exports, and that the effect is larger than tariffs. ODA flows in 2008 show that 37% of sector allocable ODA now go to aid-for-trade projects and programmes and overall aid for trade has increased to USD 41bn (2008 constant). A large increase in this period has been for transport and storage in general and for road transport in particular. Evaluation in this area will become more important. Consequently the OECD has launched a meta-evaluation for transport and storage (See Section 6 below) and many of the examples cited in this paper are related to evaluation of road projects.

#### *Results Chains*

11. Because of the "black box between outputs and impact"<sup>5</sup>, the results chain must be clearly articulated and attempts made to identify the specific impacts and how they link together. Projects should have specific objectives, and a clear idea of how these objectives will be achieved. Potential results need to be considered. For instance, the upgrading and expansion of the roadway network (i.e., primary, secondary and rural) can reduce transportation costs, improve the level of service, and provide important linkages to national, regional and international markets. In addition to market access, rural roads provide the poor with access to social services such as health and education. Given the importance of road investments within aid for trade, rigorous impact evaluations can provide lessons for future funding.

12. Cost benefit analysis can be useful in establishing criteria for undertaking specific projects, *i.e.* identifying a particular road upgrade projects. Cost-benefit analyses of roads projects typically quantify two main benefit streams:

1. Saved travel time through the ability to drive faster on better roads yielding, in turn, increases in household income through reduced prices for goods and services and increased business investment through reduced costs for inputs and firm value added
2. Reductions in wear and tear on vehicles, yielding savings in vehicle operation and maintenance expenditures

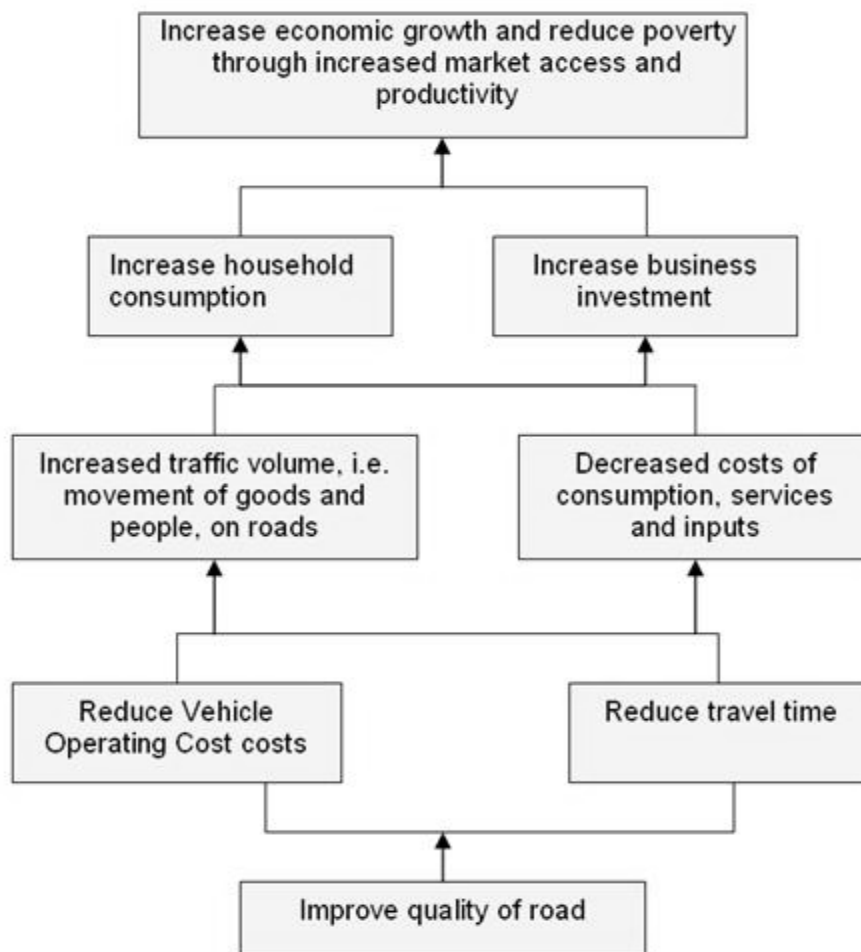
---

<sup>4</sup> If there are evaluations which are being planned by member agencies which are not included on the list, please let us know and we will update it.

<sup>5</sup> See Charles Lindahl whose presentation can be found here: <http://www.oecd.org/dataoecd/57/33/44082931.pdf>

13. A logic model (figure 1) is used to identify expected outcomes. It outlines the links in the chain from the project outcome (improved road quality) to the national level impact (growth and poverty reduction). Forming such expectations is essential in both justifying particular projects but also quantifying the expected results which can then be used to track progress and evaluate if the project succeeded in its ultimate objectives.

**Figure 1: MCC Improving Road Quality Logic Model<sup>6</sup>**

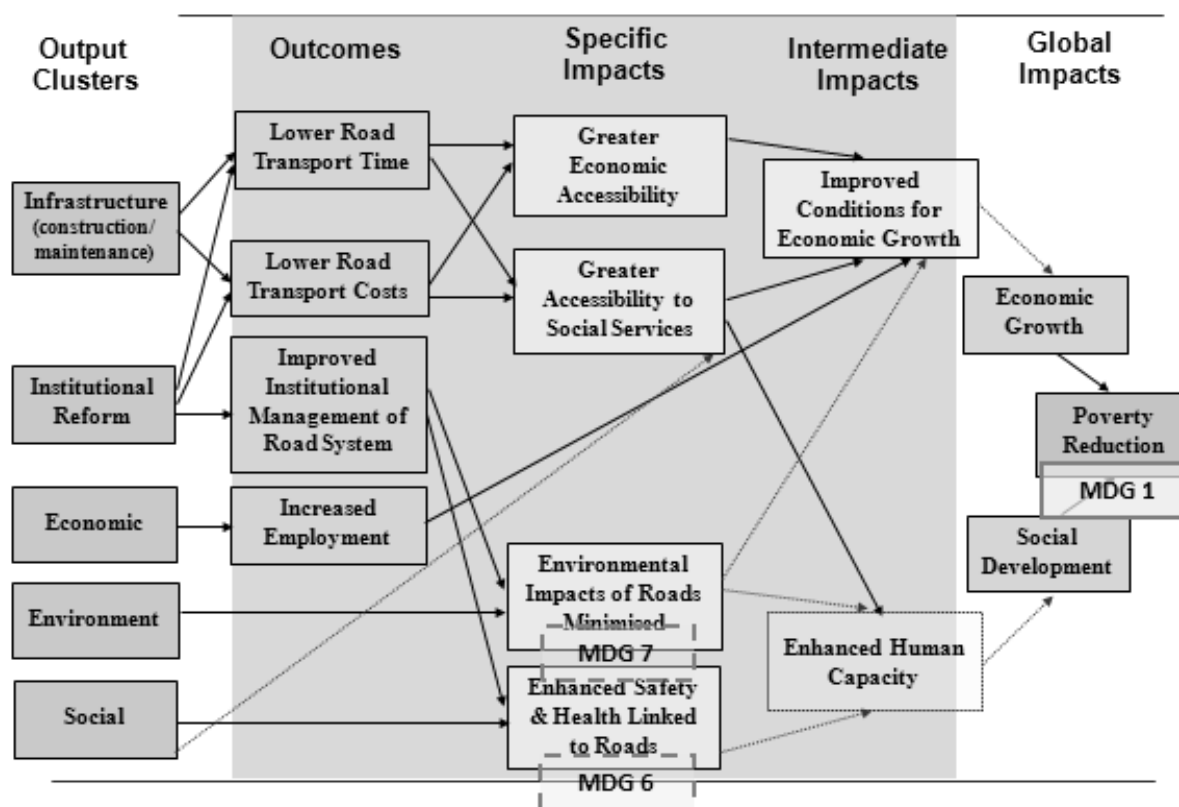


14. The Millennium Challenge Corporation (MCC) model above considers the direct economic impacts and how they link together from the output of improving the quality of roads to the overall impact of increasing growth and reducing poverty. The impacts in between are largely measurable – did the project reduce costs and increase traffic volume providing metrics against which to measure success or failure. The logic model for MCC does not articulate increasing trade as a direct outcome. However there is reference to increasing movement of goods, which indicates commerce if not necessarily international trade.

<sup>6</sup> MCC's Logic Model for Roads can be found here:  
<http://www.mcc.gov/mcc/panda/activities/impacetevaluation/ie-sectors/sector-roads.shtml>

15. The EU has developed interventions logics which are more complicated and articulate not just economic but also social and environmental impacts. They highlight impacts along the results chain and consider not just the costs and time savings but also the impacts of institutional reform, environmental consequences, enhanced safety and health linked to the roads. It then links the specific and intermediate impacts to global impacts and the contribution roads can make to economic growth and social development.

Figure 2: Intervention Logic for Roads



16. The example for roads in Figures 1 and 2 are potentially useful for thinking about other aid-for-trade interventions. However the logic models are simplified tools, not templates. An 'intervention logic' must be developed for each country based on its context and existing government strategies. At country level, the logic models are likely to be similar and developing general models to stimulate thinking about the series of impacts of aid-for-trade interventions would be helpful in establishing the evaluation questions and the data needs of the evaluation. Furthermore, there may be some benefit in providing specific logic models/intervention logics for aid-for-trade that focus on the intermediate and final trade impacts. While the models clarify different expected outcomes and impacts, impacts are not fully attributable to development interventions. Outputs are 100% attributable but attribution falls as you move along the chain of results.<sup>7</sup>

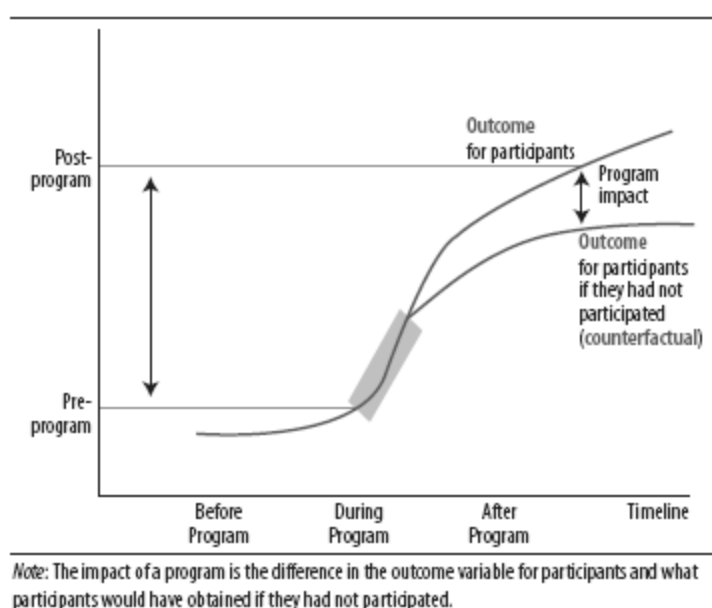
<sup>7</sup> Other donors are encouraged to discuss their approach to determining impacts and how they deal with the attribution issue.



### Impact evaluation

17. There is a growing political demand among academics, NGOs, international agencies and voters for results in development. This has stimulated new thinking about how such results could be demonstrated. Impact evaluation is one such innovation which is currently being applied in the evaluation of a range of development activities.<sup>8</sup> The approach focuses on “with and without” interventions and compares the outcomes for participants in a programme with a control group. Figure 3 sets out the timing of the programme and how the impact is determined by comparing outcomes for participants against the counterfactual. Impact evaluation typically involves a large number of units of intervention, e.g. children, households, firms, schools.<sup>9</sup>

**Figure 3: The Evaluation Problem<sup>10</sup>**



18. Identifying appropriate control groups is not easy in evaluating aid-for-trade projects. In addition application of impact evaluation to aid for trade, as noted in the previous paper, is hindered by a lack of clarity in targets and data. Without baseline information and a controlled experiment, it is difficult to attribute programme success or failure to the programmes themselves, as opposed to the environment in which they operate and the unpredictable shocks, positive or negative, that influence their effectiveness (Birdsall, 2008:528). An additional problem is that projects or policy reforms do not take place in isolation; various projects and reforms interact making identification of the actual benefits stemming from the project

<sup>8</sup> The International Initiative for Impact Evaluation provides a database which covers impact evaluations conducted in low- and middle- income countries. It summarises study findings and methodologies for both researchers, and programme managers or policy makers, who need evidence on what works, in designing and implementing impact evaluations. The database includes studies that quantify impact using a counterfactual – implicitly or explicitly – with a focus on final welfare outcomes, using qualitative, quantitative or mixed methods. The database is found here: [http://www.3ieimpact.org/database\\_of\\_impact\\_evaluations.html](http://www.3ieimpact.org/database_of_impact_evaluations.html)

<sup>9</sup> Examples of small number of observations are policy reform and many (but not all) capacity building projects. These are more amenable to qualitative approaches.

<sup>10</sup> From Fitzsimons and Vera-Hernandez (2009)

intervention precarious (For more discussion of this issue see COM/DCD/TAD(2010)/4). Nevertheless randomised trials have potential – they “revolutionised medicine in the twentieth century, and have the potential to revolutionise social policy during the twenty-first” (Duflo and Kremer, 2008:117).

19. The World Bank and others are currently investigating how to do better trade-related impact evaluation (IE). As was mentioned in the scoping note, the range of projects amenable to this type of IE techniques is probably limited. The IDB has carried out a number of ex-post evaluations of export promotion and the World Bank is working on a similar exercise for Tunisia. The World Bank is likely to pay special attention to trade facilitation. MCC have also applied this methodology to rural road upgrades.

### ***MCC Rural Road Upgrade Project***

20. As noted above, randomised impact evaluations are normally applied at the individual level e.g. education or public health. Establishing a credible comparison group for an aid-for-trade intervention is therefore difficult, but not impossible.

21. There are currently six rural road improvement evaluations being undertaken by MCC. They attempt to compare similar roads, some upgraded, some not. The roads which have not been improved act essentially as a control group from which to draw comparisons.<sup>11</sup>

22. Improved transportation infrastructure gives rise to complex economic interactions. Determining impacts at different levels as outlined in their logic model (figure 1) is hampered by a lack of data, lack of an historical track record and difficulties in measuring infrastructure services.<sup>12</sup> Household surveys are an important way of finding out whether the project had the desired effect on the most appropriate group to determine this, i.e. the road users. Surveys can help establish an appropriate baseline from which changes can be measured. A follow-up survey outlines the changes the road improvement has made.

23. They specifically look at impacts or changes at the household and firm level, observing local prices and incomes in the area and measuring how they evolve. What were growth rates of these variables before and after the project?

24. This evaluation design considers local trade impacts; international trade would be more complicated due to a number of spill-over effects but many of the expected impacts are testable – Did a road lower input prices? Did farmers switch to export crops?

### ***Wider Applicability***

25. This approach may work for upgrading rural roads but what about major national level projects like ports or airports? For these projects interim indicators are used, (For more discussion on indicators, see work on Management for Development Results and Aid for Trade). For example a port in Benin is a major hub with no comparable port, making it impossible to do impact evaluation.

26. Nevertheless there are ways of inferring impacts. To do so a lot of work needs to be undertaken in advance of the project to determine the changes that would be expected in terms of output prices and quantities of specific products traded. MCC first undertake *ex ante* analysis (due diligence, cost-benefit

---

<sup>11</sup> Road upgrades in some areas perhaps take longer to happen; perhaps they are not the most important bottleneck in the region, or otherwise low priority.

<sup>12</sup> Data collection is often the most costly element of evaluations, once staff are trained, cross cutting a sample of projects can dramatically reduce costs.

analysis) targeted on price and income effects at the household level and specific products and sectors. Did prices and incomes change more quickly after project implementation? They examine their logic model and see if it is consistent with the evidence on the ground. Did the project differentially improve interim indicators?

27. The impacts at the micro level are important to determine so as to establish the aggregate affect on the wider economy. Macro level impacts can be determined through modelling. A number of assumptions have to be made based on the changes evaluators hope to see for instance in the time to clear customs. Parameters from the evaluation can be imbedded in a model (Computable General Equilibrium (CGE) Model, Input Output models, Gravity models etc) to arrive at simulated changes at the macro level.

28. Though these method outlined in the previous section are not perfect, they are promising approaches. The potential problems are that assumptions might drive the results and developing a CGE model is extremely costly and time consuming while existing models (GTAP, GAMS) may not be suitable for the specific project.

#### **4. How should the process of evaluation change?**

29. While this work stresses the importance of evaluation, it is important that evaluation be viewed in an appropriate context. It can help provide answers but the answers are dependent on the methods, tools, and assumptions related to the counterfactual used. More constraining than methodological constraints are political economy constraints – evaluations should highlight as much failure as success, which serves the interests of few in donor agencies. Negative evaluations would subject their work to crude political criticism and limits on new resources - it is convenient and less risky politically to minimise serious evaluation work. There are also few incentives to work together with other donors or partner country representatives.

30. This section suggests that evaluation should be a collaborative endeavour and should not be constrained politically in what it can tell us i.e. independent evaluations carried out in a joint manner and where possible involving partner countries.

#### ***Evaluations should be independent***

31. There are presently few independent evaluations of aid. Many donors do have independent units within their agencies but their independence depends on the structure of the institution. Nancy Birdsall, President of the Centre for Global Development in Washington DC has called for an independent entity to undertake aid evaluations which could be easily funded by a small proportional contribution for each donor (Birdsall, 2004). She suggests that the critical ingredients for evaluation of development assistance are that it be independent, collectively agreed, and financed by a minimum set of the large multilateral and bilateral donors (Birdsall, 2004).

32. We should not expect to be able to answer all the questions that may be asked for at the political level. Political expediency can influence the selection of evaluation procedures. As mentioned already, truly independent evaluation can highlight successes but “as chance might have it, the evidence may just refuse to cooperate. Hard evidence is simply not worth the trouble, especially if eloquence and a few carefully chosen examples can carry the day” (Banerjee and He, 2008:59).

33. Evaluations are usually launched by advocates of certain sectors who want to showcase their achievement. In fact, true believers see no intrinsic value in rigorously testing the policies they are advocating – whereas there is strong support for trade leading to development results, more needs to be understood about when and where aid-for-trade interventions make this possible.

34. Independent evaluations of trade-related assistance would be a valuable global public good. We would be able to tell a lot more about what truly works and crucially what doesn't. We would learn more easily about how to improve practice and results could be widely distributed. There would also be economies of scale in that not every donor agency would need their own teams of evaluators with significant cost savings.

35. Evaluations are expensive so full use should be made of them and each new evaluation should represent a crucial contribution to the ongoing debate. Presently there is little discussion of evaluations at a global level where overall lessons should be discussed – an independent source of evaluations would help to facilitate this.

### ***Joint evaluations should be encouraged***

36. Teaming up with other donors can be demanding - it is often easier to work alone. Many donors want to know the specific contribution their tax payer money makes to the outcomes of a development assistance initiative. While joint evaluations make sense from a learning objective, it is not always compatible with accountability.

37. Joint evaluations seem particularly appropriate for aid for trade, when trying to establish the outcomes and impacts of a range of trade-related projects and programmes on overall competitiveness and trade performance. Good knowledge is needed about the disaggregated effect of different interventions. For example a port upgrade may take a long time to have impacts especially if the ultimate impacts depend on other interventions such as upgrading farm to port roads which may take more time. If customs and trade facilitation are then layered on, what is the additional effect?

38. Indeed, one of the innovative aspects of the Aid-for-Trade Initiative is the focus on sequencing a number of interventions from different donors (i.e. trade policy advice, trade development and trade related infrastructure) into a coherent framework to build trade capacities. Rather than undertaking all these interventions simultaneously, a coordinated sequencing of projects would enhance the overall outcome. Donor coordination would be needed to learn lessons on the impacts of the individual interventions and the overall impacts when all parts are completed.

### ***Partner countries should be involved in evaluations***

39. "Is this information you are gathering from us just to help you write your report or can you really be helpful to us" (Prichett, 2009). The aim of evaluation is ultimately to improve the effectiveness of development interventions. Successful evaluation cannot be a passive activity. It cannot just be about gathering information to write reports. Soliciting partner country input into the design of evaluation is therefore crucial in generating information and knowledge that is useful. There can however be practical constraints which prevent effective partner country involvement.

40. Donors often team up with partner countries on evaluations. Some donors have assisted partner-country governments in carrying out evaluation of projects. A number of multilateral and regional institutions complete project reports in consultation with partner country stakeholders (OECD, 2009:79). Partner countries participate at the WTO and should be involved in discussions about evaluating aid for trade to get their perspective. This will be one of the themes to be explored at the WTO Committee on Trade and Development Special Session on evaluation planned for July 2010.

## **5. How do we know what works at the global level?**

41. In evaluating aid for trade the focus should not only be on micro interventions and aid-for-trade projects but must also consider the broader policy considerations, which are equally, if not more,

important. When it comes to trade, economy-wide macro reforms and actions can be the real drivers of change rather than just micro interventions and projects. As such we need to be able to bring evaluation findings together to infer broader lessons.

42. You can't build a building if the blocks are not solid. Good project level work is needed before looking at macro/country level aggregates. In addition there is also a political need to be able to talk about the macro impacts of aid for trade. Therefore the project level discussion must feed into a broader discussion at the country, regional and global level. The OECD and the WTO can help to provide this global perspective, bringing evaluations and other sources of information together and drawing conclusions about broader lessons and impacts. A number of tools seem particularly appropriate to provide this perspective:

#### ***a.) Meta-evaluations***

43. Meta-evaluations (evaluation of evaluations) are designed to aggregate findings from a series of evaluations of project or programme interventions. For instance, a meta-evaluation allows us to gauge the extent to which trade is considered in evaluations of aid-for-trade related areas such as transport, energy, agriculture, forestry etc. A meta-evaluation can determine impacts at country level, provide better guidance on project implementation and design, especially how to better take account of trade objectives in trade related development programmes. To that end the OECD has launched a meta-evaluation focused on evaluations in the transport and storage sector and interventions in two countries; Vietnam and Ghana [COM/DCD/TAD(2010)3].

#### ***b.) Synthetic Evaluations***

44. Synthetic evaluations are similar to meta-evaluations but more comprehensive in their scope and use of sources. The International Initiative for Impact Evaluation has launched a series of these studies to establish the impact of development interventions. These studies go further than a standard literature review because they aim to provide an unbiased assessment of what works and why through systematic identification of relevant studies and synthesis of quantitative and qualitative evidence.<sup>13</sup> The purpose of a systematic review is to sum up the best available research on a specific question.

45. A systematic review uses transparent procedures to find, evaluate and synthesise the results of relevant research. Procedures are explicitly defined in advance, in order to ensure that the exercise is transparent and can be replicated. This practice is also designed to minimise bias. Studies included in a review are screened for quality, so that the findings of a large number of studies can be combined. Peer review is a key part of the process; qualified independent researchers control the author's methods and results. The constraint for aid for trade may be the number of studies we are aware of (plus very few actors in aid for trade would have the resources). Synthetic reviews undertaken to date typically involve the compilation of over 1900 papers from which over 50 were chosen for careful study.<sup>14</sup>

#### ***c.) Econometric Analysis***

46. There are a number of analytical research streams which are relevant to the evaluation of aid for trade. Such work can provide a broader focus on aid for trade and its impacts beyond the project level. For

<sup>13</sup> <http://www.3ieimpact.org/syntheticreviews/>

<sup>14</sup> An example can be found here:

[http://www.die-gdi.de/CMS-Homepage/openwebcms3.nsf/\(ynDK\\_contentByKey\)/MPHG-7XECFS/DFILE/White%20DIE%20WSS%20December%202009.pdf](http://www.die-gdi.de/CMS-Homepage/openwebcms3.nsf/(ynDK_contentByKey)/MPHG-7XECFS/DFILE/White%20DIE%20WSS%20December%202009.pdf)

instance, cross-country statistical analysis can examine questions about patterns of aid, trade, growth and policy settings and how these influence the overall effect of an aid for trade programme. Cross-country research can potentially inform the trade and development community in a number of ways (OECD, 2007). However there is often distrust about regression findings, particularly in the development community whereas the trade community is more accustomed to using such methods. On the other hand case study approaches favoured by the development community tend to be dismissed because findings are not usually generalisable.

47. How can we tell which interventions work best? Wilson, Helble and Mann (2009) find that a 1 percent increase in aid for trade facilitation leads to USD818m in trade. They also present a so-called ‘rate of return’. But just because trade facilitation worked well in the sample of data they used in the analysis, should we or if so why should we expect it to do as well with another group of countries (possibly at a lower stage of development)? There seems a danger here of overselling what trade facilitation can deliver. Does this make sense in development cooperation? Can we tell? If every country has different binding constraints, does it matter? Much care must be taken in interpreting and presenting econometric results. While econometric work can produce useful analysis, it should be conducted independently and where possible peer-reviewed before it is seriously considered for policy decisions.

## **6. Applicability**

48. A view shared by many in the development community is that policy based on hard evidence is simply not practical. They fear that requiring that every initiative be justified in terms of hard evidence will bias decisions in favour of what is measurable and easy to evaluate. Furthermore, evaluation does not always produce results that can justify actions – “relying exclusively on this evidence is tantamount to considered inaction” (Banerjee and He, 2008:59). The lack of evaluations and conceptual difficulties for evaluating trade outcomes may make other areas such as health more easily justifiable for donors as there is much more reliable (reassuring) evidence for those interventions.

49. Before we can say anything about the impacts of aid for trade on trade and economic growth, we need to get the micro-interventions right, we need to gather the right data, adapt our evaluation practices and work together with other donors and partner countries.

50. Time lags are also important. Many projects and programmes need time to show results. Though there is a pressure to demonstrate impacts right now, we will only be able to do this through collaborative efforts, bringing together many sources of information to get a sense of the results at a global level.

51. The issues discussed about methods raise questions about applicability. Impact evaluation should only be used where appropriate on pilot projects, valuable, representative or important interventions. It is potentially a useful tool to help fill gaps in existing knowledge. But it is only one such tool and should be used in conjunction with others.

52. Given that the methods discussed above can be contrived, or give conflicting results, how can we convince policymakers to make decisions about resource allocations based on these findings? It is certainly not the case that every policy action needs to be justified in terms of hard evidence. There are things like macro-policy that are very hard to evaluate properly. Nevertheless evidence from a limited number of micro evaluations can provide insights into macro interventions. What is certain is that we need to know a lot more. At present evaluations do not give us the sort of information we need.

## **7. Conclusions and Next Steps**

53. Aid-for-trade projects can be disaggregated and emerging methods in impact evaluation have potentially useful application. Using macro models to simulate macro impacts are also useful, though will

not be an appropriate way forward for all donors given resources and expertise constraints. Here joint evaluation work should be explored, both for pooling of expertise for also significant cost savings. Econometric approaches are also useful but early examples of this literature applied to aid for trade show myriad interpretations, data and methodological limitations and a lack of consistency in the main findings.

54. Raising awareness in the evaluation community about the trade dimension is as important as the trade community considering more evaluations and becoming accustomed to evaluation approaches, objectives and tools. No tool is perfect and the application of tools such as IE provides particular challenges but such innovations could provide essential insights at the micro level and from those we will be better able to judge if more and better aid for trade has led to improved capacity to trade, increased growth and achievement of development objectives.

55. We encourage use of the OECD Aid for Trade Community Space which serves as an important resource for sharing information and discussing evaluation concepts. This is an essential forum which the OECD can provide. Evaluation should be on ongoing discussion; we will always be looking for ways to do things better. Donors can learn a lot from each other and OECD will provide regular updates on new resources and relevant papers.

56. This paper has focused on methodology and applications with specific examples. It has also made the point that while the tools of evaluation are crucial for assessing aid for trade outcomes and impacts, there are political economy issues which can constrain evaluation findings.

57. The final version will bring the scoping note and this note together and complement these reflections with a comprehensive overview on how to evaluate aid for trade. The paper will incorporate the findings of the Meta-evaluation on Aid for Trade being undertaken. This will provide insights for future work on country case stories and studies as well as providing key insights into what aid-for-trade related evaluations currently tell us and how they could better take account of the trade dimension.

58. This work forms part of the *performance assessment framework* output result of the joint DAC-TC 2009-2010 PWB [COM/DCD/TAD(2008)7] that looks at how to better evaluate aid-for-trade projects and programmes. This work complements the work on Management for Results in aid for trade. These two work streams and the meta-evaluation will converge into a single final output, a guidance document on how to improve the performance assessment of aid for trade.

### ***Issues for discussion***

- How can we evaluate better the linkages between aid-for-trade projects and programmes and macro impacts?
- What role can the OECD play? Should OECD develop standard intervention logic frameworks considering trade impacts? Conduct Additional Meta-evaluations or Synthetic Evaluations?
- How can more joint evaluations be facilitated? How can partner countries be engaged in evaluation? What role could the WTO play?

## BIBLIOGRAPHY

Banerjee, A. V. and He, R. (2008) “Making Aid Work” in Easterly, (ed) W. *Reinventing Foreign Aid*, Cambridge: MIT Press

Birdsall, N. (2004) “Seven Deadly Sins: Reflections on Donor Failings”, Center for Global Development Working Paper Number 50. Accessible at [www.cgdev.org](http://www.cgdev.org)

Djankov, S., Freund, C. and Pham, C. (2008) “Trading on Time” *Review of Economics and Statistics*

Duflo, E. and Kremer, M. (2008) “Use of Randomization in the Evaluation of Development Effectiveness” in Easterly, (ed) W. *Reinventing Foreign Aid* Cambridge: MIT Press

Fitzsimons, E. and Vera-Hernandez, M. (2009) “A Practitioner’s Guide to Evaluating the Impacts of Labor Market Programs” World Bank Employment Policy Primer No. 12

Millennium Challenge Corporation (MCC) (2008) “Aid Effectiveness: Putting Results at the Forefront: MCC’s New Institutional Approach” by Franck S. Wiebe, Ph.D  
<http://www.mcc.gov/mcc/bm.doc/mcc-112008-paper-results.pdf>

Organisation for Economic Co-operation and Development (OECD) (2002), “*Glossary of Key Terms in Evaluation and Results Based Management*” DAC Working Party on Aid Evaluation, OECD, Paris

OECD (2006), “Trade-Related Assistance: What Do Recent Evaluations Tell Us?”, OECD Development Centre, Paris.

OECD (2007), “Evaluating Aid for Trade: Proposals for Further Work”, OECD, Paris  
<http://tcdbd.wto.org/publish/proposals.pdf>

OECD (2008), “*Evaluating Aid for Trade: Why, How and What*” Background document for session III at OECD Policy Dialogue 3-4 November 2008, Paris

OECD (2009), “Aid for trade at a Glance 2009”, OECD, Paris

Prichett, L. (2009) “Political economy of aid evaluation: How to build a sustainable *and* effective movement” Presentation to Conference “What Would the Poor Say: Debates in Aid Evaluation” Feb 6 2009, NYU

World Bank (undated), “Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank” Independent Evaluation Group, World Bank, Washington DC