Aid for Trade effectiveness: What do evaluations say?

Claire Delpeuch and Patrick Messerlin*

The Aid for Trade (AfT) initiative has been hailed as highly successful in raising the profile of trade as a tool for development. Developing countries have increasingly mainstreamed trade in their development strategies, while donors have responded by mobilising additional resources for trade-related programmes and projects (together referred to as operations). The AfT initiative however requires a better assessment of its outcomes and impacts in terms of trade performance, poverty reduction, growth and development.

In this perspective, a "meta-evaluation" of the way aid agencies evaluate and assess their trade-related operations has been conducted by the Groupe d'Economie Mondiale at Sciences Po for the Organisation for Economic Co-operation and Development (OECD), as part of the OECD work conducted in preparation for the forthcoming publication *Strengthening Accountability in Aid for Trade*.[1] The purpose of this meta-evaluation study was (i) to compare what policy makers in charge of delivering international aid would like to know about AfT results and what existing evaluations of AfT operations offer and (ii) to identify avenues for better adequacy between evaluation needs and results.

The method used

The study focused on two countries – Ghana and Vietnam – and two sectors – transport and storage.[2] It analysed 162 evaluations of AfT operations notified to the OECD DEReC database by OECD countries and international organisations.[3] We used a combination of quantitative and qualitative methods to identify the implicit interests of evaluators, as well as the gap between these interests and the information needs of AfT policy makers.

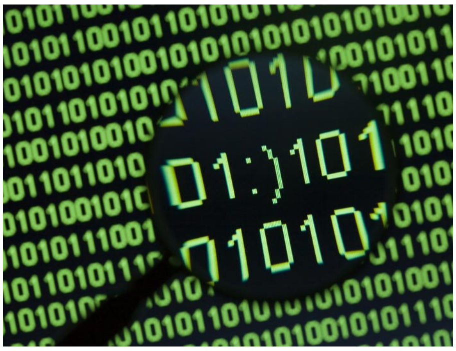
The quantitative technique was very simple: we considered the occurrence of a set of key words or expressions that could reasonably be considered crucial in evaluating AfT related operations. This method benefits from a clearly defined metric: the frequency with which key words are mentioned per 100 pages. It is however limited by the choice of key words and expressions.

The 48 key words and expressions selected in the study ranged from specific trade concerns to broad development topics (see Table 1). Indeed, while trade concerns are key because trade is a proven engine for growth, trade is not an end in itself. The ultimate goal of AfT is to help countries harness trade behind development and poverty reduction objectives – objectives which themselves have many dimensions, from income creation to income distribution to broader social concerns (such as gender equality or environmental sustainability).

The frequency with which words and expressions appear does not say much about their meaning or the context in which they are used. A critical complement was thus a systematic qualitative reading of a narrower set of the 42 evaluations identified as looking at the most trade-oriented operations. The idea was to look beyond the issues explicitly treated by the evaluators, to learn more about how each issue was evaluated and to assess the significance of the information provided.

What did evaluations look at?

The key – quite surprising – result was that evaluations of AfT operations do not say much about trade. "Trade" and "exports" were not among the most frequently used words, "imports" were mentioned even less often.



References to the World Trade Organization (WTO) and trade agreements were largely absent. In the majority of cases, the trade impact of AfT operations was not even addressed. Worse, in many instances, operations dealing with trade-related issues (such as supply-side constraints or transport and energy market issues) were never explicitly designed to achieve trade objectives.

"The key – quite surprising – result was that evaluations of AfT operations do not say much about trade."

Conversely, the evaluations referred extensively to broad development-related concepts. Most evaluations evoked the impact of operations on poverty reduction, the environment,

Table 1: List of words and expressions used in the Meta-evaluation

Trade Component	Development Component	Procedures and Techniques
Trade	infrastructure	indicator
Export	Specific regulation	performance
Import	Regulatory framework	monitoring
Trade balance	Governance	review
Comparative advantage	Supply-side constraint	impact assessment
Gains from trade	Expenditure	cost efficiency
Trade restriction	Private sector growth	cost-benefits
Tariff	Technical assistance	short term
quota	Economic growth	long term
Subsidy	Competitiveness	discount rate
Technical barriers to trade	Efficiency	counterfactual
Sanitory/Phytosanitory standards	Effectiveness	control variables
Trade facilitation	Sustainability	difference in differences
Adjustment policies	Poverty reduction	randomization
Trade assistance	gender	
Trade-related technical assistance		
WTO		
Regional trade agreements		
Preferential trade agreements		

social and gender issues as well as private sector participation. However, they did so in loose terms. For instance, gender was either mentioned in passing or measured in units that had little relevance to a project's impact on women's economic or societal situation (e.g. the number of women participating in a seminar).

Following DAC guidelines, most evaluations assessed whether project implementation deadlines were met, whether budgets were disbursed and respected as well as whether operations were in line with government's broader objectives and strategies. However, medium- to long-term impacts were often not investigated for particularly important factors such as an operation's return on investment.

In addition, the evaluations usually did not clarify the policy linkages which matter most to policy makers. For instance, potential impacts of trade policy instruments on funded projects were not examined despite the fact that high tariffs on inputs crucial to an operation would obviously limit its effectiveness – just as high tariffs imposed on import of foreign substitutes for an operation's output could artificially increase its benefits. Nor was there an evaluation of linkages with "behind the border" measures such as regulatory reforms or private sector policies.

> In addition, the evaluations usually did not clarify the policy linkages which matter most to policy makers.

How did evaluations proceed?

In most evaluations, impact assessment was solely based on the analysis of documents obtained from officials and sector stakeholders and/or on the results of field interviews. Very few evaluations relied on data collection and quantitative techniques allowing for more robust performance assessments.

Such practices are problematic. First, the use of qualitative methods was not always satisfying: it was rarely explained how the available documents were analysed and how results were derived. Second, it is difficult to assess outcomes exclusively through use of qualitative methods. Qualitative methods often leave open the problem of attribution, that is, the measurement of the part of the observed changes that results from AfT operation versus from exogenous factors. In some evaluations, economic analysis was used to predict the impact of operations, and to judge their relevance. While this approach does not address the causality problem, it does improve the quality of the *a priori* evaluation of an operation.

To be fair, all these criticisms need to be put into perspective. Failure to refer to specific trade results can be explained, at least in part, by the absence of stated trade-related objectives in the operations' initial mandates. In most cases, little *ex ante* economic analysis of the operations had been undertaken, resulting in a lack of quantifiable definitions of objectives and of baseline data that would have made possible a comparison of the *ex ante* and *ex post* situations. Moreover, it is often difficult to assign macroeconomic outcomes and impacts to individual AfT operations given the complex array of variables influencing outcomes. Finally, chosen objectives sometimes conflict. For instance, in one evaluation, support to the agricultural sector was said to have had a positive pro-poor impact through successful targeting of poor recipients although it failed to promote growth or export diversification, raising serious questions about its sustainability.

Conclusion

The study demonstrated that, to date, in many AfT evaluations, causal links between operations and performance are not established – be performance defined in terms of trade, poverty reduction or intermediate objectives. Consequently, their conclusions give little insight into whether AfT works and why.

The study also showed that there were few clear links between donors and the issues assessed or the methods used. In other words, there were few systematic "leaders" in terms of evaluations of a given topic for a given country and/or sector. This suggests that a broader sample of evaluations would not have produced substantially different results.

In a time of harsh budgetary constraints, continued increases in AfT contributions require a better understanding of what works and how. One option to achieve this understanding is to use sophisticated techniques that enable thorough evaluations of the impacts of operations. Such projects are currently being undertaken, notably by researchers of the World Bank. Yet, as such techniques are expensive and time consuming; they can be used only for a carefully chosen subset of AfT operations. Hence, there is also a need to identify simple guidelines that would already improve substantially evaluations, while requiring minimal time, skills and financial resources. The meta-evaluation study suggests that every operation should include an *ex ante* assessment of the situation defining quantifiable objectives and providing the information (including data) necessary to measure whether objectives were met *ex post*. In this context, it recommends that evaluators should answer systematically a thorough sequence of questions that would provide a framework for assessing operations' impact in a more thorough way.[4]

Authors

Claire Delpeuch is an agricultural economist in the Trade and Agriculture Directorate, OECD, working mainly in the areas of food security, agricultural policy, trade and development. She is also associate researcher at the Groupe d'Economie Mondiale at Sciences Po.

Patrick Messerlin is professor of economics at Sciences Po, and director, Groupe d'Economie Mondiale at Sciences Po (GEM). He specializes in international trade policy and regulatory reforms. His current research deals with WTO issues, EC commercial policy, services liberalisation and the associated "Better regulations" initiatives

***Other contributors to this article include :**

Marie-Agnès Jouanjean is a research fellow at the Groupe d'Economie Mondiale at Sciences-Po and at LICOS, K.U.Leuven. She works mainly on agricultural trade and poverty dynamics.

Alexandre Le Vernoy is a consultant and research fellow at the Groupe d'Economie at Sciences-Po. His research topics include trade, environment and natural resources.

Thomas Orliac is a consultant and research fellow at the Groupe d'Economie Mondiale, working mainly on trade facilitation.

This article is based on a full paper by Delpeuch Claire, Marie-Agnès Jouanjean, Alexandre Le Vernoy, Patrick Messerlin and Thomas Orliac (2011) "Aid for Trade: A Meta-evaluation". The study can be found on the OECD website and on the GEM website.

Notes

1 Delpeuch Claire, Marie-Agnès Jouanjean, Alexandre Le Vernoy, Patrick Messerlin and Thomas Orliac (2011) "Aid for Trade: A Meta-evaluation". The study can be found on the OECD website and on the GEM website.

2 Ghana was selected because it is one of the largest recipients of aid for trade in Africa and has seen its trade expand significantly since 2002. Vietnam was selected because it has served as a model of trade opening for a number of countries; it has enthusiastically embraced the AfT agenda; it has successfully achieved export-led growth; and it became a WTO member fairly recently, in 2007. The transport and storage sectors were selected because they have received significant amounts of AfT and have therefore been subject to a large number of relevant programme and project evaluations.

3 The DAC (Development Assistance Committee of the OECD) Evaluation Resource Centre (DEReC) can be accessed here

4 This set of questions could for instance be based on the evaluation guidelines of the Australian Office of Best Practice Regulation that identify a series off steps in the realisation of the impact assessment.