AFT's Impact: How Should We Evaluate?

DE MELO, Jaime, CADOT, Olivier, NEWFARMER, Richard

Abstract

Did the AFT initiative make a difference? It did help to mainstream trade in donor strategies and increase AFT flows. Yet, the initiative was remarkably poor in terms of evaluation mechanisms or even simple guidelines on how to conduct evaluations. We review here key considerations on how to better assess AFT's impact.

Reference

AFT’s Impact: How Should We Evaluate?

How should we evaluate?

August 2013


Did the AFT initiative make a difference? It did help to mainstream trade in donor strategies and increase AFT flows. Yet, the initiative was remarkably poor in terms of evaluation mechanisms or even simple guidelines on how to conduct evaluations. We review here key considerations on how to better assess AFT’s impact.

The Aid-for-Trade (Aft) initiative will soon be ten years old, a good time to look back and discuss whether it genuinely made a difference or not. The initiative largely resulted from a convergence of views, the trade community realizing that the costly commitments made by developing countries in the Uruguay Round called for trade-adjustment assistance—especially for low-income ones—and the wider development community embracing at least partly the notion that trade was, in itself, an effective poverty-alleviation mechanism.

Following the 2005 Hong Kong Ministerial, the WTO Task Force on Aid for Trade summarized its objectives as “assisting developing countries to increase exports of goods and services, to integrate into the multilateral trading system, and to benefit from liberalized trade and increased market access.” AFT would “enhance growth prospects and reduce poverty in developing countries, as well as complement multilateral trade reforms and distribute the global benefits more equitably across and within developing countries”.

Did the AFT initiative make a difference? It certainly helped to mainstream trade in donor strategies and to reverse the decline in the share of Overseas Development Assistance (ODA) going to trade-related activities. Before 2005, industrial-country efforts to help the integration of low-income countries into the world trading system largely relied on trade preferences; while the largest part of ODA targeted education, health and direct poverty-reduction programs. The initiative seems to have reversed the trend, with AFT flows rising from US$17.8 billion in 2005 to US$27.5 billion in 2011—although the growth leveled off in recent years—and boosting the share of trade-related activities in total commitments from 30% in 2005 to 35% in 2010. The trend reversal is most visible for multilateral donors, who were also the ones having reduced most vigorously their trade-related activities since their peak in the days of structural adjustment (Figure 1).

Figure 1. Share of AFT disbursements in total foreign aid, 2002-2010

However, with tighter budget constraints in donor countries and louder demands for accountability across the board, AFT is likely to face in coming years a “results test” going beyond increased aid flows. Yet, the initiative was remarkably poor in terms of evaluation mechanisms or even simple guidelines on how to conduct evaluations. So far, the quest for accountability has produced a digest of a large collection of projects and case stories—voluntarily supplied and thus heavily selected—feeding into meta-analyses built around word counts.
(OECD 2011). Whatever quantitative evidence there is on AfT’s impact comes from evaluation exercises carried out in academia, without explicit linkages with or funding from the initiative’s governance structures.

It has been known since the pioneering work of Amjadi and Yeats (1995) and Limão and Venables (2000) that trade costs are very large for low-income countries, especially in Africa, explaining a large chunk of the continent’s poor trade performance. From that perspective, the AfT initiative’s emphasis on infrastructure investment (63% of commitments, of which 35% was for roads and 16% for rail) is justified empirically. However, justification is not evaluation; in fact, recent research—see e.g. Carrère and de Melo (2009), Novy (2012) or Arvis et al. (2013)—suggests that trade costs have declined less rapidly for low-income countries, the prime target of AfT, than for others, reinforcing their insulation from world trade.

Beyond aggregate evidence on trade costs, direct attempts to evaluate ex-post the impact of AfT have produced an ambiguous picture. Recent work attempting to identify the impact of targeted programs to reinforce productive capacities (e.g. Brenton and von Uexkühl, 2009; Cali and te Velde, 2011; Ferro, Portugal and Wilson, 2012) has produced weak evidence. The problem with ex-post evaluations based on publicly-available data is that the identification of AfT’s impacts is vulnerable to so many confounding influences that even with the most creative econometric techniques, it is unlikely that sufficiently clear-cut evidence could emerge to respond to hard-headed demands for results-for-money.

In order to credibly claim impact, AfT needs to shift its evaluation paradigm—a paradigm that has emerged largely spontaneously so far—to one that focuses on “cutting the length” of the causation chain. That is, instead of incidentally trying to assess whether larger aid flows reduce trade costs or raise exports, on average, over large panels of heterogeneous countries, it should move to a paradigm where projects are explicitly designed, at the outset, for direct impact evaluation, by careful construction of treatment and control groups. The key requirement here is not necessarily randomization per se—whether or not randomization is the alpha and omega of impact evaluation is the subject of an intense debate in academia—but more modestly to carry out baseline and follow-up surveys on sufficiently large samples, including project beneficiaries and non-beneficiaries, prior to the intervention and after it.

Here is where difficulties start. First, by construction, treatment effects capture only effects that are internalized by the beneficiaries. But then, why shouldn’t they pay for them? Subsidised interventions (most aid for trade takes the form of grants or concessional loans) should be justified by some sort of market failure such as non-appropriability of the gains, as funds have an opportunity cost. But if gains are not appropriable, they won’t show up in a treatment-effects test. Thus, the absence of estimated treatment effects suffers from a basic ambiguity; it could be that the program was ineffective, in which case it should be discontinued, but it could also be that its effects spread to the control group, in which case it should be continued (it could also be that the test does not have sufficient power to reject the null, a sample-size problem). In plain English, impact evaluation can be a key piece in the monitoring-evaluation nexus, but it should be interpreted cautiously.

Second, situations of ‘clinical’ policy interventions in trade are rather rare. Targeted programs such as technical assistance for export promotion could be amenable to randomized control trials or other forms of impact evaluation, but the more numerous non-targeted reforms like customs reforms, port improvements or other institutional improvements are less easily amenable to the usual methods (although sometimes it is still possible to go down from the intervention level, say a border post, to the firm or transaction level, as in Volpe and Graziano, 2012).

Third, implementation faces two types of constraints, i.e. incentives and costs. As for incentives, project manager buy-in would be facilitated if impact evaluation could be fully decoupled from their evaluation, but no organisation could commit to that without facing a time-consistency problem. As to costs, bottom estimates for an evaluation are around US$300,000 . For large-scale social or health projects, typically this will be only a few
percentage points of programme cost. But trade-related projects are much smaller, so containing evaluation costs to 5% of project costs (requiring project cost above US$6 million) will put the majority of aid-for-trade projects outside the range of feasibility. Cadot et al. (2012) estimate a median commitment size of US$700,000 (aggregated over all donors) for trade policy and regulations. In conclusion, randomised control trials face an uphill road in trade-related assistance but quasi-experimental methods relying on existing data from customs and industrial surveys provides a second-best alternative.

The way forward: Using benchmarking to identify programme effects

For both hard and soft infrastructure, causal links from policy intervention to export performance are strongly suggested by theory but non-trivial and often elusive to estimate empirically. Cross-country evaluations will continue to be needed because they are the safest route in terms of ‘external validity’, in spite of their limitations in terms of ‘internal validity’ (ability to establish causality from intervention to effects). In order to generalise the use of impact evaluation methods in trade-related interventions, given the typically small size of such projects, what is needed is to make it practically feasible in terms of design (project and evaluation using quasi-experimental methods), incentives (impact evaluation results should be decoupled from individual performance evaluation), and resources (get government buy-in to release confidential data). Governments will be more willing to relinquish semi-confidential data to researchers if they understand the value of the results generated.

Oliver Codot is Professor of International Economics at the University of Lausanne, and Research Fellow at FERDI and Centre for Economic Policy Research CEPR. Jaime de Melo is Professor at the University of Geneva and Research Fellow at the Fondation pour les études et recherches sur le développement international (FERDI). Richard Newfarmer is Country Director for Rwanda, South Sudan and Uganda at the International Growth Centre.

References


Carrère, C and J de Melo (2009) “The Distance Puzzle Resides in Poor Countries”, VoxEU.org, 10 November.


OECD (2011) “Strengthening Accountability in Aid for Trade” OECD, Paris


This article was published in Great Insights Volume 2, Issue 5 (July-August 2013)