FONDATION POUR LES ÉTUDES ET RECHERCHES SUR LE DÉVELOPPEMENT INTERNATIONAL

LA FERDI EST UNE FONDATION RECONNUE D'UTILITÉ PUBLIQUE.
ELLE MET EN ŒUVRE AVEC L'IDDRI L'INITIATIVE POUR LE DÉVELOPPEMENT ET LA GOUVERNANCE MONDIALE (IDGM).
ELLE COORDONNE LE LABEX IDGM+ QUI L'ASSOCIE AU CERDI ET À L'IDDRI. CETTE PUBLICATION A BÉNÉFICIÉ D'UNE AIDE DE L'ÉTAT FRANÇAIS GÉRÉE PAR L'ANR AU TITRE DU PROGRAMME « INVESTISSEMENTS D'AVENIR » PORTANT LA RÉFÉRENCE « ANR-10-LABX-14-01 »

Aid for Trade: Can it be Evaluated?

Olivier Cadot
Jaime de Melo

Olivier Cadot is professor of International Economics and director of the Institute of Applied Economics at the University of Lausanne and Senior Fellow at Ferdi. He is also Research Fellow of the Center for Economic Policy Research in London and an Associate Scholar of Cerdi and the Cepremap.

Jaime de Melo is professor at the University of Geneva since 1993 and Senior Fellow at Ferdi since 2011. His research focuses on links between regionalism and multilateralism, trade policies and migratory policies, trade and environment.

Following the recent surge in Aid-for-trade (AFT) and the likelihood that it will decline in the future, pressure to evaluate its effectiveness is increasing. This brief summarizes some of the main lessons from a recent FERDI-ITC-WB workshop. While recognizing that a multiplicity of approaches are needed to learn what works and what does not, this brief argues that the emphasis on reducing trade costs is well placed in spite of the lack of firm evidence linking AFT flows to measures of trade costs because trade volumes are consistently found to be responsive to variations in trade costs. However, relying on cross-country studies to detect AFT ('hard' or 'soft') effects is inconclusive. Impact evaluation (IE) is not a panacea, but is a credible alternative.

policy brief
58
January 2013
note brève

• • • / • • •

···/··· But IE faces three difficulties: (i) the treatment may spread to the control group in which case it is not discernible; (ii) situations of "clinical interventions" in trade are rare; (iii) incentives and costs are a hurdle in implementation. The brief concludes that the way ahead is to use benchmarking (from existing data sets) to identify program effects.

The demand for Accountability is on the Rise…

At their annual conference in Hong Kong in 2005, WTO trade ministers called for expansion of Aid for Trade (AFT) to help "developing countries, particularly LDCs, to build the supply side capacity and trade-related infrastructure that they need to implement and benefit from the WTO Agreements and more broadly to expand trade". This expansion of AFT reflected a recognition that internal constraints--trade-related infrastructure (ports, roads and transport or 'hard' infrastructure) and trade-related institutions (customs, standard agencies, and policies and regulations affecting trade, or 'soft' infrastructure) were becoming more important than traditional barriers at the border. Those have been drastically reduced by the 'negative agenda' built around the reduction of the traditional (tariffs and quotas) external barriers.

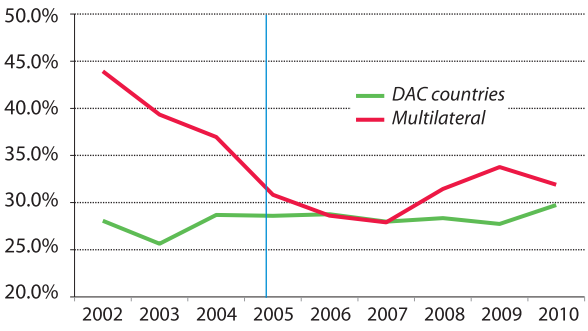
In brief, AFT was to reduce trade costs. A WTO AFT task force was set up in 2006 to implement this 'positive agenda' to enhance competitiveness. Multiple goals were adopted[1], but clear guidelines on how to conduct evaluations were largely absent even though pressure for greater accountability was mounting. Down the road, after three biennial reviews organized around the OECD-WTO task force, what we have is a discussion of approaches and methods, leading to a 'managing for Development Results' (MfDR) approach along a 'results chain' to enable evidence-based evaluation based on firm definitions and clear objectives. The quest to improve accountability has produced case studies and a digest of a large collection of projects and case stories--many voluntarily supplied and thus heavily selected--feeding into meta-analyses built around word-counting (OECD, 2011).

…but aggregate data shows no discernible aggregate effects in spite of rising AFT volumes

Start with AFT volumes, the yardstick keenly followed by negotiators, in particular from developing countries. By the commitment measure, the 2005 initiative has indeed been highly successful, reversing the long-term decline in the share of trade-related assistance in ODA. The trade-related share rises from 30 percent in 2005 to 35 percent in 2010, boosting annual commitments (from $25 billion in 2005 to over $45 billion in 2010). Tracking disbursements (figure 1) shows that the "big push" to the AFT agenda came mostly from multilateral agencies rather than from bilateral Aid from DAC countries. There was no noticeable change in the allocation categories, with hard infrastructure taking the lion's share (63%, of which 35% was for roads and 16% for rail), against 29% for behind-the-border policies (a hodge-podge of projects including sectorally-targeted ones), 6% for technical assistance on trade policy—arguably the most "T"-related in the AFT—and 2% for trade facilitation. Of course, these numbers reflect not just donor priorities but also intrinsic cost differences, as building a bridge is costlier than computerizing a border post.

1. The WTO task force listed in order: increasing trade, diversifying exports, maximizing linkages with the rest of the economy, increasing adjustment capacities, regional integration, an contributing to inclusive growth and poverty reduction.

Figure 1. Aid for Trade Disbursements, 2002-2010 (share in Total ODA)

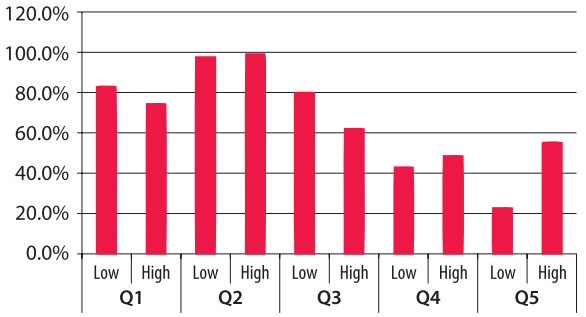


AFT's potential to increase exports by reducing trade costs has been the centerpiece of this agenda. Figure 2 checks whether a simple correlation between lagged (to allow for delays) AFT disbursements and export growth is visible to the naked eye. It splits the set of AFT recipients within quintiles of the (baseline) export/capita distribution by the median into two cohorts, "low recipients" and "high recipients", based on average 2000-2005 receipts. Thus, Q1 is the worst-performing quintile in the baseline period, Q2 is the second-worst, and so on. Results are striking: only in the top two quintiles can one see a positive export-growth differential between high- and low-recipients (Panel a). AFT could, however, have an indirect effect on export performance by working primarily through improved logistics markets. Panel b checks for this by carrying out the same exercise for the time to export, with similarly disappointing results.

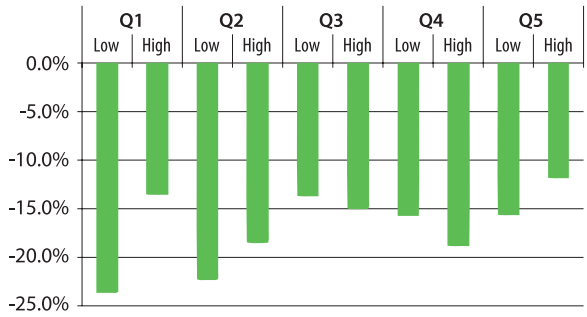
Inspection of project write-ups has also proved somewhat inconclusive. Word counts from several hundred project write-ups revealed that what matters most for policy makers (terms like "imports", "exports" or "regulatory reform") were rarely mentioned. The review also highlighted that project evaluators often lacked the baseline data against which to measure progress so that quantitative indicators were absent from all, but a handful of write-ups (e.g. data on investments in infrastructure even though 80% of AFT in low-incomes is assigned to infrastructure development).

Figure 2. Export growth and time to export vs. lagged AFT, by quintile of the export per capita distribution

(a) Export growth (five-year cumulative)



(b) Time to export



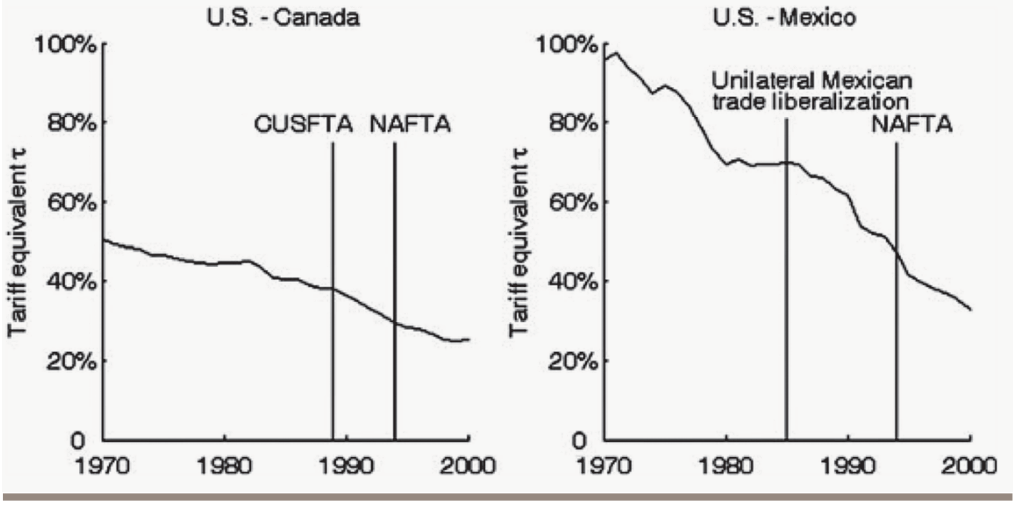
Source: Cadot et al. (2012).

...but the evidence does support the reduction of trade costs as a means to increase trade

In spite of the ambiguous prima-facie evidence on impact, both theory and empirics suggest that the emphasis on reducing trade costs is not misplaced. Indeed, a whole family of well-established models of international trade gives paramount importance to trade costs in the determination of trade: controlling for other determinants, countries that are more closed to trade because of high bilateral trade costs relative to

domestic costs, are expected to trade less. This is the 'gravity equation', from which one can compute aggregate bilateral trade costs from observed bilateral trade data. As confirmed in the quasi-experimental case of NAFTA bilateral trade costs fall the most most (and hence bilateral trade increases the most) during the period of reduction in trade barriers.

FIGURE 1
The U.S. Relative Bilateral Trade Cost Measure with Canada and Mexico



Source: Novy (2012)

Where should AFT be directed: hard or soft?

Unfortunately, clean experiments like NAFTA are rare, especially among heavy recipients of AFT. There are many determinants of aggregate trade costs, they fluctuate frequently, and they are hard to measure, let alone to disentangle from volume effects. Should AFT be directed towards improving 'hard' infrastructure (roads, ports, railways)? Or should it target 'soft' infrastructure (the provision of backbone services, regulation, competition)?

A review of the evidence in a large number of studies (Cadot et al. 2012) shows some progress: controlling for a host of other determinants to trade, all components of trade costs, however measured, reduce the volume of trade. But when it comes to suggesting which reductions in trade costs have the largest impact, no pattern emerges. For example, in their pioneering study, Limaõ and Venables (2000) showed that a few variables capturing hard infrastructure and geography explained the large variation in shipping costs across destinations. Variation in their proxy for 'hard' infrastructure (a composite index of roads, rail and telephone lines) contributed 50 percent of the variation in container freight rates across destinations, whereas distance only contributed 10 percent. By contrast, after a careful study of trucking corridors across Africa, Teravaninthorn and Raballand (2008) concluded that, most often, road rehabilitation was not the bottleneck on improved trucking performance—the problem was rather in regulatory frameworks. In other words, donor policy dialogue with African governments to improve competition in service provision was as important, if not more, than building roads.

Compounding somewhat ambiguous results on what should be the priority between hard and soft, evidence on the impact of AFT flows on trade costs is also largely inconclusive, partly because only a small part of the literature has taken the step of looking for impact, and partly because identification is inherently difficult given the many confounding influences. Cali and te Velde (2011) found that AFT had more impact when targeting hard infrastructure than when targeting particular sectors. Vijil and Wagner(2012) similarly found some impact for aid to infrastructure, while Ferro, Portugal and Wilson (2011) found that aid to upstream service sectors had a significant effect on downstream manufacturing performance (using IO tables as the linkage).

Impact evaluation: No panacea, but a credible alternative

Impact evaluation (IE) has spread to development economics partly in response to the identification difficulties faced by cross-country econometrics, because it typically provides a much sharper way of identifying "treatment effects" from aid interventions to performance, provided that interventions have some "clinical" (individual-level or firm-level) dimension. Essentially, IE identifies the impact of projects by comparing the performance evolution of treated entities with that of a control group. The key data requirement is a baseline survey administered on a sufficiently large sample including both beneficiaries and non-beneficiaries prior to the intervention as well as a follow-up survey.

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? Subsidized interventions (most AFT takes the form of grants or concessionnal 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, IE 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 IE, but the more numerous non-targeted reforms like customs reforms, port improvements or other institutional improvements are less easily amenable to the usual IE 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: incentives and costs. As for incentives, project manager buy-in would be facilitated if IE could be fully decoupled from their evaluation, but no organization could commit to that without facing a time-consistency problem. As to costs, bottom estimates for an IE are around $300'000 . For large-scale social or health projects, typically this will be only a few percentage points of program cost. But trade-related projects are much smaller, so containing IE costs to 5 percent of project costs (requiring project cost above $ 6 million) will put the majority of AFT projects outside the range of feasible IE. For instance, rough calculations from the OECD CRS data base by Cadot et al. (2012) estimate a median commitment size of $700'000 (aggregated over all donors) for trade policy and regulations.

Given the well-known limitations of impact

evaluations and experiments in terms of "external validity" (their ability to generate results that carry over to different settings), it should come as no surprise that randomized control trials face an uphill road in trade-related assistance.

Does this mean that IE is a lost cause in the evaluation of AFT? Not so fast. Short of full-fledged randomized-control trials, quasi-experimental methods relying on existing data from customs and industrial surveys provides a second-best alternative. The recent literature evaluating export-promotion agencies (see Volpe, 2011) provides a striking example of the ability of those methods to generate policy-relevant results.

The way forward: Using benchmarking to identify program 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). Improvements in data, both in terms of quality and accessibility of existing data will lead to more generalizable results As to IE methods, given the typically small size of trade-related projects, In order to generalize the use of IE in trade-related interventions, what is needed is to make it practically feasible in terms of design (project and evaluation), incentives, and resources. In terms of design, the message of our brief overview of methods is that there is substantial scope for adapting methods to the particular context of trade interventions, especially with quasi-experimental approaches. In terms of incentives, we argued that if the decision to launch an IE and budget for it out of project resources is left to project managers, there is an agency problem. Part of the problem is the potential for IE to bring bad news. Thus, IE results should be decoupled from individual performance evaluation, but promises to keep a firewall between the two are unlikely to be time-consistent. Thus, the decision to launch IEs should be taken upstream of project management. One solution might be to set up an independent IE center for AFT projects. However, ultimately government buy-in would be a crucial ingredient, and it would be unlikely with a complete separation of IE from project management. Cooperation across donors and close dialogue with the government is necessary for buy-in.

Finally, and perhaps most importantly, adopting IE as routine practice in AFT projects requires the "evaluation community" to work on reducing IE costs. Although experienced IE practitioners like to warn newcomers against "doing IE on a shoestring", the currently very high cost of IEs acts as a powerful deterrent. In trade policy, there should be scope for better use of existing statistics and, crucially, for more dialogue with governments to ensure the availability of firm-level statistics that have been collected under different auspices. That is where the issues of cost and buy-in converge: Governments will be more willing to relinquish semi-confidential data to researchers if they understand the value of the results generated.

References

- **Cadot, O.**, **A. Fernandes**, **J. Gourdon**, and **A. Mattoo** (eds.) *Where to Spend the Next Million? Applying Impact Evaluation to Trade Assistance*; London/Washington, DC: CEPR/ The World Bank , http://www.voxeu.org/sites/default/files/file/next_million_WB.pdf

- **Cadot, O., A. Fernandes**, **J. Gourdon**, **A. Mattoo**, J. de Melo « Evaluation in AFT : From Case-study Counting to Measuring », London/ Washington, D.C. CEPR/World Bank http://www.ferdi.fr/uploads/sfCmsNews/html/177/DT_evaluation%20in%20AFT.pdf

- **Cali, M. and D. te Velde** (2011). "Does Aid for Trade Really Improve Trade Performance?" *World Development*, 39(5), 725-40.

- **FERDI-ITC-WB** (2012) "Aid for Trade: What Have we Learnt? Which Way Ahead?" http://www.ferdi.fr/evenements/Aide+au+commerce.html workshop, Geneva, December 6,2012.

- **Ferro, E.**, **A. Portugal-Perez**, and **J. Wilson** (2012). "Aid to the Services Sector: Does it Affect Manufacturing Exports?", mimeo, paper presented at the FERDI-ITC-WB workshop.

- **Limao**, **Nuno** and **Anthony J. Venables** (2000), "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade", *World Bank Economic Review*, 15(3) : 451–479

- **Novy, D.** (2012) "Gravity Redux: Measuring International Trade Costs with Panel Data", *Economic Inquiry*

- **OECD** (2011) " Strenghtening Accountability in Aid for Trade" OECD, Paris http://www.oecd.org/aidfortrade/48365377.pdf

- **Teravaninthorn**, **Supee** and **Gael Raballand** (2008), *Transport Prices and Costs in Africa: A Review of the Main International Corridors*, Washington, DC: World Bank.

- **Vijil, Mariana**, and **L. Wagner** (2012), "Does Aid for Trade enhance export performance? Investigating the infrastructure channel"; *World Economy*, 35(7), 838-68

- **Volpe, Christian** (2011), "Assessing the Impact of Trade Promotion in Latin America"; in O. Cadot et al. eds., pp. 39-81.

- **Volpe, Christian**, and **Alejandro Graziano** (2012), "Customs as Doorkeepers: what are their effects on international trade?"; mimeo, paper presented at the FERDI-ITC-WB workshop.

FERDI

Créée en 2003, la **Fondation pour les études et recherches sur le développement international** vise à favoriser la compréhension du développement économique international et des politiques qui l'influencent.

Contact
www.ferdi.fr
contact@ferdi.fr
+33 (0)4 73 17 75 30

policy brief
58
January
2013
note brève