FONDATION POUR LES ÉTUDES ET RECHERCHES SUR LE DÉVELOPPEMENT INTERNATIONAL

INTERNATIONAL BANK FOR RECONSTRUCTION AND DEVELOPMENT
WORLD BANK

International Trade Centre

Evaluation in AFT:

From Case-study Counting to Measuring

By

Olivier Cadot **Ana Fernandes** **Julien Gourdon**

Aaditya Mattoo **Jaime de Melo**

Paper prepared for the workshop :
"Aid for Trade: What Have we Learnt? Which way Ahead?"
6 December 2012, International Conference Centre, Geneva

Revised version: January 2013

1. Introduction

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 or standard agencies), policies and regulations that discourage trade or 'soft' infrastructure were becoming more important than the barriers to trade resulting from policy barriers at the border which 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, but clear guidelines on how to conduct evaluations were largely absent.[2] Evaluation has progressed slowly from accountability (making sure that infrastructure has been built) to outcomes (has performance improved, but no agreement as to the main yardsticks to be used to measure outcomes). Progress has also been slowed by donors (multilateral, bilateral and NGOs) using different evaluation frameworks, by lack of information, and by context-specificity.

So far, the OECD-WTO task force and three biennial reviews have produced a useful discussion of approaches and methods and a digest of a large collection of projects and case stories--many voluntarily supplied--feeding into meta-analyses built around word counting. For example, the meta evaluation of 162 projects in Ghana and Vietnam (not all with a trade emphasis) revealed that what matters most for policy makers (terms like "imports", "exports" or "regulatory reform") were rarely mentioned. It also highlighted that project evaluators often lacked the baseline data against which to measure progress. A review of case-stories, rich in project details, indicated a lack of quantitative indicators and revealed large gaps in emphasis (only 3 of 269 reported on aid for trade adjustment and few reported on investments in infrastructure even though 80% of AFT in low-incomes is assigned to infrastructure development).[3]

To intervene more effectively, donors and recipients need more rigorous assessments of AFT projects. As shown by the case studies reported in Cadot et al. (2011) and those discussed here, many of the interventions described in these case studies, be they technical assistance, export promotion, or programs targeted at women entrepreneurs/traders, could be evaluated rigorously, provided that impact evaluation is part of program design from the outset – and that donors and

[2] According to OECD (2011), the AFT agenda has been classified under six categories: (i) trade policy and regulation; (ii) trade development; (iii) trade-related infrastructure; (iv) building productive capacity; (v) trade-related adjustment; (vi) other trade-related needs. According to OECD (2011), 80% of donors use the DAC principles for evaluating programmes and projects (see box 1 for the 5 criteria—relevance(suitable), effectiveness(achieves objectives), efficiency(least-cost approach), impact, sustainability

[3] The meta-evaluation was carried out by Messerlin et al. ( ) and the case stories by Folletti and Newfarmer (2010). These are reported in OECD (2009) and OECD (2011).

beneficiaries are willing to commit the resources necessary to undertake the work. But impact evaluations, while indispensable to tackle the attribution problem, are resource-intensive and raise doubts when it comes to generalizations to other environments. Data collection across countries that are amenable to cross-country economy-wide (macro) evaluations are therefore a useful complement to the (micro) impact evaluations. When controlling for other intervening factors, the macro studies help detect the regularities across countries. Reviewing the results from these macro studies establishes a few stylized facts. For example, regardless of measures and estimation methods, improvements in trade facilitation have been found to be more important determinants of export performance than policy-imposed barriers at the border.

We now have in hand a large portfolio of macro cross-country estimates of trade costs and their impact on aggregate export performance. With a few exceptions, the missing link is an estimate of AFT on performance (volume of trade, diversification, or growth), in part because the aid flows are not categorized along the categories identified in the AFT objectives and, in any case, have multiple impacts, some of which are hard to measure—a far-reaching regulatory reform may not necessitate large technical assistance.[4] While pointing in the direction that AFT result in more favorable indices of trade costs that are then associated with better trade performance, the results from these macro studies do not address the attribution problem because the lack of a convincing counterfactual. This is where the growing number of impact evaluations is helpful because of their ability of netting out confounding influences.

The survey follows this script. Section 2 classifies the components that enter into trade costs and discusses how trade costs influence the volume of trade. Section 3 reviews the studies on trade costs where discussion is organized around the relative importance of 'hard' vs. 'soft' infrastructure in aggregate trade costs and trade volumes. The following sections turn to the links between AFT disbursements and performance. Section 4 summarizes the results from economy-wide studies linking AFT flows to outcome measures. Section 5 then turns to a discussion of impact evaluations, the different methods and the circumstances, when they are likely to be applicable to AFT. Section 6 concludes.

2. Trade Costs: Classification and Measurement

Following the third Global review, evaluation of AFT is to adopt a 'results chains' approach where activities give rise to project outputs (better quality hard infrastructure or better indicator values for soft infrastructure) which contribute to final outcomes, measured by increased trade which presumably will lead to higher growth (and other objectives, e.g. reduced poverty). Here we are concerned with data needs to enable evidence-based evaluation at the macro level. Reduction in trade costs is the first stage of the evaluation process.[5] For ease of exposition and clarity, we

[4] WTO Task force listed as objectives "increasing trade, diversifying exports, maximizing linkages with the rest of the economy, increasing adjustment capacities, regional integration, and contributing to inclusive growth and poverty reduction" (OECD (2011), 17))

[5] DAC criteria for Evaluating Programmes and Projects accepted by over 80% of donors are the following:

- Relevance (suitability, e.g. increase exports and diversify the export base)

associate AFT outputs with a reduction in trade costs and associate final outcomes with increases in exports (greater diversification could be another final outcome). Section 2.1 catalogues the elements that enter into trade costs and the links between project outputs (as captured by indicators of trade costs) and final outcomes (as measured by indicators of aggregate trade). Section 2.2 shows how the workhorse gravity model provides the required link between aggregate trade costs and exports volumes. Section 2.3 discusses how gravity-based studies have given evidence of the various components of the trade costs that are then reviewed in section 3.

2.1 AFT and Trade: the Channels

Figure 1 indicates the links between AFT activities and quantifiable outcomes. AFT is at the bottom, directed to several components of trade costs that determine trade flows (an approximate average breakdown of AFT expenditures over 2006-10 is given next to the corresponding arrows). The thickness of the solid lines is suggestive of the relative importance of the linkages that has emerged across studies, or in the case of AFT flows, of the relative sized of the flows.[6] Solid arrows indicate the direction of causality that is generally recognized in the studies and the dashed arrows indicate attempts at handling two-way causality. For example, many studies find that indicators of geography and hard infrastructure capture a larger portion of the variance in trade costs than indicators of border and behind-the-border policies. Most studies also find that differences in the values of indicators of the quality of hard-infrastructure contribute more towards accounting for differences in trade costs than differences in geography.

Note two boxes from which no arrows emerge. First are political commitments that are essential to the ownership and success of AFT projects--- key to follow-through and close monitoring of AFT activities at the country-level---that are not quantifiable in a meaningful way. Second are macroeconomic policies which have been shown to be an important influence on exports but that are not considered in the trade cost evaluation literature, except occasionally through time fixed-effects in panel studies.[7]

- Effective (achieves objectives , e.g. reduce protection, improve road network, reform customs)
- Efficiency (least cost approach, e.g. maintenance of existing road network rather than expansion)
- Impact (appropriate indicators, e.g. trade costs and trade volumes)
- Sustainability (benefits continue after donor funding ends)

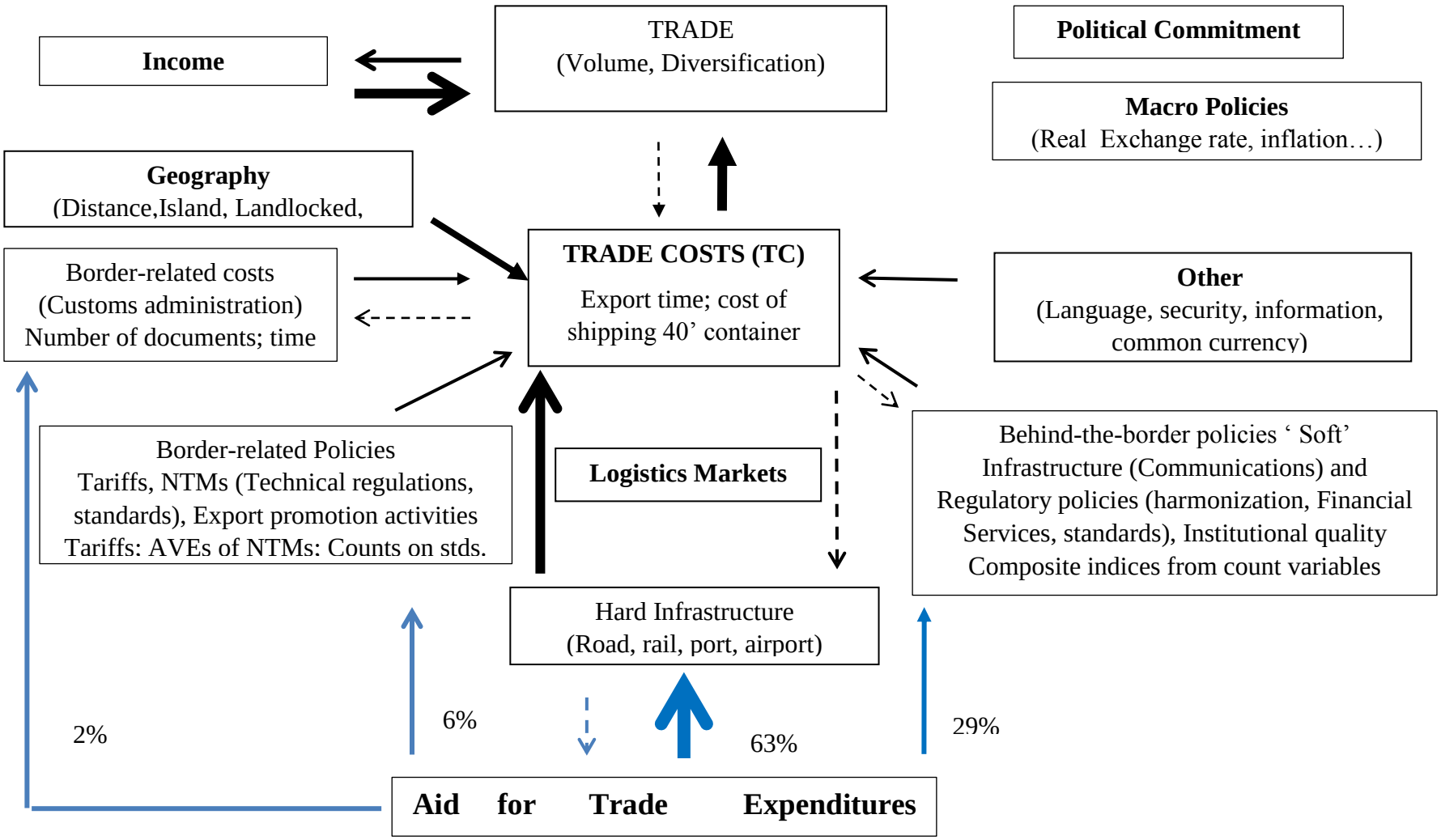
In a survey administered to evaluators of projects, 50% of respondents answered that assigning trade outcomes to the programme presented them with the most difficulty way ahead of the other criteria listed above (see OECD (2011), figure 1.1.)

[6] Motivation and allocation of AFT across countries is discussed in section 4.1

[7] Freund and Pierola (2012) show that export surges are strongly correlated with sustained real exchange rate depreciations and that this is particularly so for the extensive (new products) margin, perhaps because market failures affect tradables more than non-tradables . Given the importance attached to the diversification of exports in AFT objectives, this should be taken into account.

Figure 1: From AFT to Trade: Hard and soft Linkages[a]

(thickness indicates strength of linkages)



[a]AFT Disbursements (CRS source) : Average percentage shares (2006-2010) exclude AFT for Energy and Productive S

AFT expenditures in the bottom of the table intervene to reduce trade costs through improved infrastructure ('hard' and 'soft') and technical assistance to improve policies at the border and behind. Average AFT disbursement shares targeted to these objectives (according to a modified OECD classification—see annex table A2) over the period 2006-10 are indicated next to the arrows. They show that the lion's share of AFT (63%) goes to infrastructure.

Four aspects of these linkages deserve attention. First, as indicated in the top left corner, there is a strong two-way causality between trade and income not yet resolved but not discussed in the AFT literature where the concern is the causality running from AFT to trade income or growth (via improved infrastructure and policies).[8] Second, the main channel investigated is the causality running from hard infrastructure and geography to trade costs after controlling if possible for other determinants (language, trade policies, various behind-the-border policies). Third, except for a few studies, the causality running in the other direction is not taken into account, thereby contaminating the estimates.[9] Fourth, logistics 'markets' in the middle of the figure usually do not figure explicitly in the evaluation so that one does not know the underlying cause of high freight rates along a trip remain unknown.[10]

2.2 The Primacy of Trade Costs

That international trade costs are very large has long been established. What is important is how they are measured and how they are disaggregated.[11] Virtually all evaluations of the determinants of aggregate trade costs are based on some version of the ubiquitous gravity equation. Indeed, Novy

[8] Frankel and Romer (1997) and others have concentrated on this two-way causality. As to the link between aid and growth, after reviewing 97 studies, Ducougliagos and Paldam (2011) conclude that the literature on aid efficiency has failed to establish a causal link from aid to growth. These results are echoed in Bourguigon and Sundberfg (2007, p. 317) who conclude "if a dollar of aid produces little discernible change, was the objective ill-defined, the service delivery inefficient, bureaucratic measures inadequate, or was money diverted?".

[9] Often researchers are careful not to read causality in their results. For example, Brenton and Von Uexkull (2009), conclude (p.250) that "while questions remain on the issue of causality, there can be little doubt that, on average export development programs have coincided with or predated stronger export performance in the targeted commodities" . In contrast, reporting on these results, the OECD report on strengthening accountability states "that the authors concluded that exports increase as a result of export development programs and projects" (OECS (2011), p. 22). In effect, the issue of whether or not fast growing exports attracted funding for the 86 export development projects they studied remains.

[10] Logistics costs extend beyond transport costs to include transaction costs, financial and nonfinancial costs and would be measured by the price paid at the point of consumption.

[11] According to the rough estimates of Anderson and van Wincoop (2004), even for trade between the US and Canada, two adjacent countries under free trade that share a common language and a largely common heritage, they have been estimated at an ad-valorem of 44%.The breakdown of 44% is : 8% on policy, 7% on language, 14% on currency, 6% on information, 3% on security. See Anderson and Van Wincoop (2004). These high costs to cross the Canada-US border are confirmed by price differences in products defined at the bar code level sold by large retail stores on both sides the border where wholesale and retail prices of the same product differ by 24% across the border (see Gopinath, Gourrinchas and Li (2011).

(2011) shows that a quasi-identical trade cost function obtains from three popular trade models.[12] For familiarity's sake take the well-known Anderson van Wincoop (AvW) model. In equilibrium, bilateral trade flows are given by

$$X_{ij} = \left(\frac{Y_i Y_j}{Y^W}\right)\left(\frac{TC_{ij}}{\Pi_i P_j}\right)^{-\eta}; \quad \eta > 0 \tag{1}$$

where TC refers to the trade costs described in Figure 1 and the denominator is the endogenously determined 'multilateral resistance terms'. Since bilateral trade barriers affect internal and international trade, letting (t_{ii}, x_{ii}) and (t_{jj}, x_{jj}) denote intra-national trade barriers and intra-national trade flows, countries with a more closed economy will have higher multilateral terms. This can be seen from (2) which expresses the product of multilateral resistance terms in terms of trade flows:

$$\Pi_i P_i = \left(\frac{x_{ii} / y_i}{y_i / y^w}\right)^{-\eta} t_{ii} \tag{2}$$

Take two countries *i* and *j* of the same size and the same domestic trade costs. Then from (2), the more closed country (say *i* if $x_{ii} > x_{jj}$) will have a higher multilateral resistance term. By measuring bilateral trade costs relative to domestic trade costs, Novy obtains the following measure of ad-valorem bilateral trade costs in terms of observables:

$$\tau_{ij} = \left(\frac{x_{ii} / x_{jj}}{x_{ij} / x_{ji}}\right)^{\zeta} - 1; \quad \zeta > 0 \tag{3}$$

The expression shows that if bilateral trade flows increase relative to domestic trade flows, this must have happened because of a reduction in bilateral trade costs. This expression becomes operational after constructing intra-national trade flows obtained by taking services out of GDP. Thus, (3) provides an ad-valorem estimate of total bilateral trade costs (tariffs, language barriers, currency barriers, the equivalent of NTMs, etc..) without having to specify a functional form for trade costs.

[12] Novy (2011) shows that the gravity equation comes out of three family of micro-founded trade models. In the best-known Anderson and van Wincoop (2003) model, production is exogenous and each country consumes all goods by trading a single good with its partners—trade is driven by the love of variety. Eaton and Kortum (2002) concentrate on the supply side in a Ricardian model where each firm's productivity is drawn from a distribution leading each country to specialize in the lowest cost good it can produce—trade is driven by relative productivities. Finally, in the heterogeneous firm models of Chaney (2008) and Ottaviano and Melitz (2008), firm productivities are also drawn from a distribution. In Chaney, each firm faces fixed costs of exporting, while in Melitz and Ottaviano firms face sunk costs of entry and firms have a comparative advantage in technology. Novy shows that when the ratio of domestic to bilateral trade flows is considered, all the country-specific variables that determine bilateral trade drop out leading to the bilateral trade cost equation given here. Because it is best known, we give the expression for the AvW model rather than for one the others that model the supply side (the only differences across models come from the exact determinants of η --tastes in AvW and parameters of the distributions in the other models).

From a sample of 13 OECD countries covering 1970 to 2000, Novy estimates that average trade costs of Canada [Korea] fell from 131% [246%] to 101% [146%]. He also estimates that trade costs between the US and its NAFTA partners dropped more rapidly during the period of NAFTA implementation, showing the benefits of market integration. Comfortingly, he also shows that his constructed measure is correlated along expected lines (distance increases trade costs and adjacency reduces them). Finally, conditional on the validity of the gravity model (trade-income causality left out), he decomposes the growth in trade into income growth, decline in relative bilateral trade costs and relative multilateral resistance. These contributions vary across countries and partners. For US trade with Mexico, the percentage contributions in parenthesis are: income growth (54.8%); decline in relative trade costs (57.4%); decline in relative multilateral resistance (-12.2%), the negative estimate capturing the decline in trade costs of Mexico with other partners which partially offsets the decline in bilateral trade costs. Taking out general equilibrium effects, about one third of changes in trade are attributable to a relative decline in trade costs, largely justifying directing AFT activities towards reducing trade costs and evaluating it in terms of progress on trade cost indicators.

Given the distinguished performance of the gravity model in explaining trade flows, inspection of aggregate trade costs and their evolution over time across countries should be useful in the evaluation of AFT activities. For example, in Novy's OECD sample, the average rate of decline in trade costs over the period was 35% with the highest rate of decline in Korea at 55%. Such estimates for a larger sample of developing countries might serve as indicators of where efforts at reducing trade costs would have the highest payoffs. In any case, focusing on how to reduce trade costs is important to promote trade.

2.3 Measuring Trade Cost Components and Disentangling their Impact

The first step in the analysis consists in getting a handle on the measurement of trade costs. Transport costs are captured by a trade cost function TC with the arguments listed in (4):

$$
\begin{aligned}
TC_{ij} &= T(V_{ij}, \tau_{ij}, b_i, b_j, \lambda_{ij}) \\
V_{ij} &= (D_B, D_L^i, D_L^j, D_{ISL}^i, D_{ISL}^j, DIST_{ij}) \\
\tau_{ij} &= (1 + t_{ij}^{ave}) \\
b_i &= b^i(INFR_i, IC_i; AFT_i) \\
b_j &= b^j(INFR_j, IC_j; AFT_j)
\end{aligned}
\qquad (4)
$$

In (4), trade costs are a function of various combinations of the elements identified in figure 1. First are the characteristics that relate to the journey between origin and destination, V_{ij} (modal choice, dummy variables capturing invariant geographic factors---common border, whether one or both of the countries are landlocked or an island, and distance). Next is the ad-valorem equivalent of trade barriers τ_{ij}, often approximated by the MFN tariff of the country of origin, but sometimes including the ad-valorem equivalent (AVE) of NTMs, t_{ij}^{ave} . Then comes a vector of characteristics of each

trading partner, b_i and b_j, which captures the quality of infrastructure. Infrastructure includes 'hard' infrastructure, $INFR_i$ (indicators of the quality of roads, etc…) and 'soft' infrastructure (e.g. an investment climate indicator like the costs of trading across borders or an indicator of institutional quality or governance $\left(IC_i, IC_j\right)$). Finally, all unobservable variables are grouped under λ_{ij}.

Since the pioneering work of Limão and Venables (2001), international trade costs, have often been measured by the cost of shipping a standard container between the two trading partners which is a shortcut for . Alternatively, an aggregate measure of a country's cif/fob ratio is a more encompassing, but less reliable, measure of trade costs.[13]

AFT enters (4) via two channels: (i) By reducing trade barriers and improve indicators of 'soft' infrastructure (tariffs, NTM measures, harmonize standards) through technical assistance activities; and (ii) by improving the quality of hard infrastructure. Unfortunately, only the last channel can be credibly measured, hence "the fear [by the development community] that requiring that every initiative be justified in this way [basing policies on hard evidence] will bias decisions on what is measurable and easy to evaluate" OECD (2011, p. 33). As discussed in section 4, so far only three studies have explored the link between AFT and exports.

To move ahead, data are needed on the hard and soft components of trade costs. In the last decade, the World Bank and others (e.g. WEF) have compiled data on trade facilitation, the investment climate, information on delays, and on administrative procedures. These have been compiled and elaborated into several composite indices such as the Logistics Performance Index (LPI) database, Doing Business (DB) Surveys, and World Bank Enterprise Surveys. Typically, gravity-based estimates either use the raw data from those surveys (Chen et al. 2006 or Portugal and Wilson 2008) or a ranking index (Hoekman and Nicita 2008). Information on non-tariff measures, available on Trains or in the World Bank enterprise survey, have been used as dummy variables (Chen et al. 2006 or Helbe et al 2009) or as an AVE (Hoekman and Nicita 2008, Portugal and Wilson 2008). Several studies apply principal component analysis (PCA) or factor analysis (FA) on different indicators to generate a composite index (see e.g. Helbe et al (2009) for their transparency index, Wilson et al (2003), Balcin and Edwards (2008), or Francois and Machin (2006) for their infrastructure and institution index).

All these indicators are composite for which one has difficulty relating improvements in the values of these indices to their underlying constituents since the weighing schemes are not related to any technological relation and, the indices are often proxies rather than direct measures. Finally some studies conduct their own survey to extract specific information. These might come from freight forwarding companies who provide data on time for exports of a standard container (Djankov et al 2010) or Duval (2006)) or from experts on trade facilitation measures under negotiations at WTO.[14]

[13] Measurement error and holes in the data are rife. See Limão and Venables and Hummels and Lugovskyy (2006).

[14] Djankov et al. collected their data from a survey administered to 345 freight forwarders in 126 countries for 2005 consists of the time it takes for a 20' standard container carrying three categories of standard goods (textiles and

With trade cost estimates in hand, the second step relates trade flows to trade costs (only in a few studies discussed in section 3 are trade costs related to AFT flows as suggested above). This is typically done with the gravity model. Substituting (4) into(1), rearranging and linearizing, an estimation by OLS using fixed effects could take the form :[15]

$$\begin{aligned}
&\log X_{ij} = C + F_i + F_j + \phi[\log TC_{ij}] + \log \varepsilon_{ij} \\
&C = -\log Y^W \\
&F_i = \log Y_i - \log \Pi_i + b_i \\
&F_j = \log Y_j - \log P_j + b_j \\
&\log TC_{ij} = \sum\nolimits_k a_k V_{ij}^k + c\tau_{ij}
\end{aligned} \tag{6}$$

In this expression all terms that are partner-specific (such as INFR and IC and the unobserved multilateral resistance terms) are confounded as they are lumped in the country fixed effects $\left(F_i, F_j\right)$. This precludes estimating the effects $\left(b_{i(j)}\right)$ as they only vary by exporter (constant across importers) or vice versa. Thus, since the model is one of bilateral trade (which has the advantage of giving the degrees of freedom that would otherwise be absent in cross-section), one can only estimate the effects of variables that vary bilaterally and indexed over k, such as distance, common language, common border, ad-valorem equivalent of bilateral trade barriers. [16] This considerably limits the usefulness the of the gravity model which in effect is extended to a problem that is not bilateral to buy degrees of freedom. Moving to panel estimation can help, but for many indicators of trade costs, the data are only available on a cross-section basis or, at best, for a few years. Estimation must also address the possibility of endogeneity of the regressors.

Finally, implementing (6) requires addressing the large number of zero observations in the data. Two methods are used: Use of the Poisson Pseudo Maximum likelihood estimator proposed by Santos Silva and Tenreyro (2006) handles heteroskedasticity coming from the fact that the error term is in logarithms making it impossible to properly account for zero observations by OLS. However, this approach does not deal with the problem that the omitted variable (zero trade flows) which accounts for positive trade flows could be correlated with one of the regressors, in which case a Heckman correction model could be applied provided that one can obtain a 'good' instrument for the first-stage regression.

---

apparel, and coffee, tea and spices). The freight-forwarding companies reported the time it takes to carry out four stages of exporting procedures (pre-shipment activities, inland carriage, port handling, customs and technical control). Their estimates used several approaches to deal with the endogeneity of trade costs to the volume of trade. The number of freight forwards in each country is limited and unlikely to be representative of trade costs (see Volpe and Graziano (2012)).

[15] In cross-section, the multilateral resistance terms identified by the theory can be handled by fixed effect techniques (i.e. by grouping terms for exporters and importers) so that the estimated equation captures the trade elasticities to the cost variables.

[16] There are fixes to this problem, for example by taking variables that vary only by exporter or importer and transform them artificially so that they vary bilaterally. They come at a cost as one can no longer link results to an underlying model. Random effects estimation is not a solution either as it requires strong assumptions about the unobserved heterogeneity in the data (i.e.. the multilateral resistance terms should be normally distributed)..

3 Infrastructure, Facilitation, and Trade

3.1 The 'Hard' Side: Roads and Ports

> "The plan was to carry 1,600 crates of Guinness and other drinks from the factory in Douala where they were brewed to Bertoua, a small town in Cameroon's south-eastern rainforest. According to a rather optimistic schedule, it should have taken 20 hours, including an overnight rest. It took four days. When the truck arrived, it was carrying only two-thirds of its original load. …we were stopped at road-blocks 47 times. … Our road was rendered impassable by rain three times, causing delays of up to four hours. The Cameroonian government has tried to grapple with the problem of rain eroding roads by erecting a series of barriers … that stop heavy trucks from passing while it is pouring. … Early on the second evening … we met a [locked] rain barrier in the middle of the forest. It was dark, and the man with the key was not there. … he returned shortly before midnight. The hold-up was irritating, but in the end made no difference. Early the next morning, a driver coming in the opposite direction told us that the bridge ahead had collapsed, so we had to turn back."[17]

Beyond longstanding interest in the question of how transport costs--especially maritime costs that account for 80% of world trade—have evolved (see Moneta, 1959, or more recently Hummels, 1998), attention has turned to the constraints on LDC exports created by poor infrastructure. This emphasis arose from the observation of Africa' s poor export performance in the late 1990s in spite of wide-ranging structural adjustment reforms.[18] For instance, in an early study, Amjadi and Yeats (1995) found that over 40% of the export earnings of some of Africa's landlocked countries were absorbed by freight and insurance payments, with a continent-wide average of 15% (against 5.8% for all developing countries).

A breakthrough came with Limão and Venables' pioneering study (Limão and Venables 2000), where they introduced shipping company quotes for a 40ft container carrying standard good as an alternative measure to cif/fob price comparisons as a measure of trade costs. They approximated 'hard' infrastructure by a composite index of roads, rail and telephone lines which they showed contributed 50 percent of the variation in container rates across destinations while distance only contributed 10 percent of the total variation.[19] Their results were far reaching. Not only did they show the superiority of this measure to cif/price comparisons in estimating bilateral trade patterns, but they showed that a few variables capturing infrastructure and geography along the journey described in equation (4) went a long way towards explaining the large variation in shipping costs from Baltimore to 64 destination cities across the world. In addition to confirming the high costs of

[17] Anecdote reported in *The Economist*, 2002, cited in Buys et al. (2006).

[18] Frankel (1997) found that "under-trading" was particularly acute in the case of intra-regional trade. Classic papers by Collier (1995) and Collier and Gunning (1997) attributed Africa's under-trading to the disastrous policies including (inter alia) protectionism, currency overvaluation and export monopolies, adopted roughly between the mid-70s and mid-90s. However, Foroutan and Pritchett (1993), Coe and Hoffmaister (1998), and Rodrik (1998) argued that size, income and other gravity determinants largely explained Africa's low trade volumes.

[19] As noted by Limaõ and Venables who were the first to introduce a composite index of infrastructure, taking a linear combination of these components assumes that these inputs are perfect substitutes. Bundled up with capital and labor in a Cobb-Douglas function gives a cost function for transport costs.

being landlocked, they detected additional costs to overland distance (1'000 km of overland distance added on average $1'380 to container freight costs, against only $190 by sea) for landlocked countries compounded by border delays, uncertainty, higher insurance costs, and charges by transit countries. They also were the first to show that a composite index of infrastructure was an important contributing factor to overall costs. Thus, an improvement from the 75th percentile to the median for their infrastructure index would be equivalent to a distance reduction of 3'466km by sea or 419km overland. Finally, they also showed that this estimated transport cost estimate performed very well in a standard gravity equation, estimating that a 10% reduction in trade costs increased trade by 30%.[20] Their key finding was that 'hard' infrastructure accounted for nearly half of the transport cost penalty borne by intra-SSA trade. Applied to Africa, Limaõ and Venables' estimates suggested that poor infrastructure overexplained the underperformance of the continent's trade. Coulibaly and Fontagné (2006) confirmed Limão and Venables' results on aggregate and disaggregated trade flows in West Africa, predicting that if all roads were paved in the region, trade would almost treble.

The policy implications of this body of work were clear: It provided intellectual support to a return of the "big-push" view, according to which donors should build roads and ports in order to unlock Africa's trade and, by implication, its growth. Indeed, Buys, Deichmann and Wheeler (2006) built on Coulibaly and Fontagné's results to explore the returns on a pan-African program of road infrastructure development. They identified inter-city corridors for road investment using spatial-network software and used gravity coefficients to estimate the investment program's impact on trade. Finally, they used World Bank data on the cost of road improvement and rehabilitation ($127'000/km for the median project) to perform a cost-benefit analysis. After several ad-hoc adjustments for local variations in costs (e.g. due to rainfall) and efficiency (e.g. due to bad governance), the results were stunning: The payback horizon was barely over one year, with $254 billion of additional trade generated over the project's estimated lifetime at a cost of about $32 billion. A similar exercise, in which trade was disaggregated by sector, was performed by Shepherd and Wilson (2006) for the ECA (Europe and Central Asia) region. Interestingly, trade elasticities to infrastructure were substantially lower than to other variables (e.g. tariffs); however, the scope for road improvements were such that it remained an extremely favorable policy proposition. For instance, a complete upgrading of the road infrastructure in Hungary, Romania and Albania (at about $227'000/km) would generate an "on-impact" (annual) trade increase of over $35 billion for a one-time cost of $3 billion.

The extremely high rates of return on road investments identified in the trade literature were substantially above, but broadly consistent with, high rate-of-return estimates found in the macro growth literature—typically over 200%—which put road investments on top of other types of

[20] If $\hat{\phi} = \partial \ln x_{ij} / \partial TC_{ij}$ is the trade elasticity to a trade-cost variable (say, distance or infrastructure in *i*) and $\hat{\delta} = \partial \tau_{ij} / \partial x_{ij}$ is the CIF/FOB elasticity to the same variable, the estimated effect of CIF/FOB margins on trade is $\hat{\phi} / \hat{\delta}$. Alternatively, they used the determinants of CIF/FOB margins as instruments in a two-step approach Both approaches yielded elasticities around three, i.e. very strong inhibiting effects of transport costs on trade.

infrastructure investments such as telecommunications or energy (see Estache 2004 and references therein).[21]

Thus, after almost two decades of multilateral donor emphasis on structural adjustment and policy reform, by the mid-2000s empirical research was suggesting that the pendulum should swing back toward (infrastructure) capital accumulation.[22]

3.2 The 'Soft' Side: Facilitation, Regulation and Competition

While evidence accumulated on the strong effect of infrastructure on trade costs, whether the right policy response was a "big push" in infrastructure investment was questioned by Shantayanan Devarajan, the World Bank's Chief Economist for Africa, in his foreword to Teravalinthorn and Raballand (2008):

> "One of the few things that African policy makers, development partners, civil society, and policy researchers agree on is that Africa has a serious infrastructure deficit. […]While everyone agrees on the problem, there are different approaches to a solution. One view is that, if Africa has an infrastructure deficit, the solution is to plug that deficit by investing in infrastructure—build new roads, power plants, and irrigation canals. Another is to identify the causes of Africa's infrastructure deficit and address them directly. For, if the problem is policy or institutional failures that prevent infrastructure from being productive—irrational power tariffs, weak regulations, inadequate operations, and poor maintenance—then simply building new infrastructure without addressing these problems will not improve the situation. Africa will still have an infrastructure deficit—but with higher debt." (pp. xi-xii)

This preoccupation reflects a new awareness of the importance of the 'logistics markets' in figure 1. Responding to concerns of this type, in parallel to the literature on infrastructure and trade, another strand of the literature had developed in the 1990s on the regulation of infrastructure and transportation services. An early study by Bennathan, Escobar and Panagakos (1989) drew lessons from Chile's deregulation of its shipping industry, which until 1979 had been regulated by a cargo reservation system dating back from the 1950s, increasingly at odds with the Pinochet government's pro-market orientation. Like a number of other Chilean policy experiments, this one proved a success: Not only did freight rates rise more slowly in Chile than in countries with cargo reservation systems, but by the late 1980s the national carriers' share of national shipping was unchanged, suggesting that the old system was merely protecting rents.

[21] These too-good-to-be-true rates of return were reminiscent of the "Aschauer debate" on infrastructure and growth (see Estache and Fay 2007 and references therein for an overview). An internal evaluation of World Bank infrastructure projects over 1999-2003 produced an economic rate of return of 43%, by all means a respectable rate but nowhere near the miracles suggested by the literature (see Estache 2004). However, the ranking of rates of return also put road investments on top, suggesting the same lending priorities.

[22] Meanwhile, the attention of development economists had shifted from policies (endogenous variables in the political-economy literature) to institutions, with a literature setting up "horse races" between institutions and geography as explanatory factors for trade performance (see e.g. Redding and Venables 2004 or Francois and Manchin 2006).

Since maritime transport still accounts for 80% of world transport, it is important to explore how generalizable the Chilean experience is. Fink et al. (2002) and Clark et al (2004) explored the impact of efficiency on shipping costs to the US. Fink et al. regressed freight-rate data for U.S. seaborne imports on the existence of maritime cartels (so-called "shipping conferences") as well as various restrictive regulations applying to shipping (cargo reservation schemes) and port operations. They found evidence that cartels substantially pushed up freight rates (by about a third) but the evidence on policy restrictions was inconclusive. Among the problems with both estimates is that they relied on single point in time which did not allow them to control for heterogeneity across ports and time-invariant omitted variables. More importantly perhaps, their indicator of port efficiency was drawn from interviews.[23]

As usual, better data led to more trustworthy estimates. Drawing on reliable US data on bilateral import charges at the HS-6 level over the period 1991-2003, Bloningen and Wilson (2008) regressed for each product import charges on all relevant characteristics except changes in product composition. After controlling for all other factors affecting charges, their port fixed effects provided an efficiency ranking of US and foreign ports. Overall, they estimate that a 10 percent increase in port efficiency increases trade between a country-pair by 3.2 percent, or alternatively a change in port efficiency from the 75% percentile to the 25% percentile leads to a 5 percent increase in trade.[24]

Further progress came from studies digging deeper into cartel behavior, long known to be prevalent among "shipping conferences". Inspired by the observation that Caribbean and Central American countries trade far less than would be predicted by the gravity model (Guatemala's exports of manufactures to Caribbean partners are far less than 1%, yet they are close and have easy access to each other by sea), Wilmsmeier and Hoffman (2008) analyze freight rates charged by one major liner shipping company on 189 routes in the region. Their estimates show that distance is trumped by the number of liner shipping companies providing services between pairs of countries, a result that would likely also carry over to Sub-Saharan Africa where transshipments are frequent.[25]

Again focusing on US ports and maritime traffic to Latin America, Hummels, Luggovsky and Skiba (2009) estimate the market power of shipping companies by using the cross-product variation of

[23] This measure also used by Clark et al (2004) was taken from the Global Competitiveness Report where survey-respondents were asked to rank countries rather than ports on a 1 to 7 scale.

[24] They obtained precise estimates by regressing import charges on weight, value, distance, the percentage of shipment in containers, a measure of trade imbalance, and fixed-effects that controlled for all time-invariant omitted variables (observed and unobserved). They estimated that a 1o percent increase in distance increased freight costs by 1.3-2.1 percent, that imbalanced trade raised costs but not by much and that containerization reduced import charges. In their sample, the only African port (Durban) has port charges 15 percent above those of Rotterdam. They estimate a slow improvement for foreign ports towards Rotterdam at 1.4% per year over the 12 year period. Their estimates are lower-bound estimates insofar as they do not take into account that increased efficiency would bring trade in new products

[25] Their model of liner shipping freight rates has the following variables: transshipment vs. direct services; the number of competing carriers; UNCTAD's liner shipping connectivity index; transit time; port infrastructure endowment in the importing and exporting countries. The model accounts for three fifth of the variance of liner shipping freight rates across the Caribbean.

tariffs to identify unobserved market power.[26] They estimate that eliminating market power in shipping would boost trade volumes by 5.9% for the US and 15.2% for Latin America. Furthermore, high tariffs on trade give market power to shippers: a 1% increase in tariffs leads to a 1-2% increase in shipping prices per kilo.

Turning to road transport, Teravaninthorn and Raballand (2008) showed that trucking deregulation in Rwanda after the civil war had effects similar to those of shipping deregulation in Chile: nominal rates dropped by 30% and the domestic trucking fleet expanded instead of shrinking. By contrast, countries like Malawi where domestic truckers were protected by restrictive entry regulations, ended up essentially penalizing farmers—a common policy outcome in Africa. They also highlighted the deleterious effects of cartels and regulations through "freight bureaus" on Central African corridors where freight rates per ton/km were about 80% more and truck-utilization rates 40% less than on East African corridors.[27] Throughout West Africa, they found that bilateral agreements, queuing systems and quotas stifled competition. Even on the most competitive trucking corridors of East Africa, anticompetitive regulations abounded, with e.g. Kenya prohibiting international transit trucks on the Mombasa-Kigali corridor from taking domestic freight on the return trip, forcing them to cover 1'700km empty.[28] Their conclusion was in striking contrast with those reached in papers discussed in the previous section:

> […] because of the high cost of road improvement and the relatively old fleets, rehabilitation on hundreds of kilometers of road would not be economically justified if traffic were less than 200 trucks per day. Below such traffic levels, rehabilitation probably should take place only when the road is in poor or very poor condition (*and only if the benefit from VOCs reduction were passed on to the final user of transport services*). (p. 83—italics ours)

For donors, the implications of Teravaninthorn and Raballand's work were starkly different from those of previous pieces of empirical research on infrastructure: rather than build more roads, they should pursue policy dialogue with African governments to improve regulatory frameworks and ensure competition in service provision. Should that be taken to mean that governments should step back, deregulate and merely ensure fair and open competition? Not so fast if one digs deeper and resorts to impact evaluation (see section 4).

What do we learn in terms of trade-off between reforms aiming at reducing barriers at the border and those aimed at trade facilitation more generally? Most studies find that trade facilitation

[26] When tariffs are high, the share of freight costs in consumer prices is lower, and so is the price elasticity of demand perceived by the shipping lines, which will, if they have market power, induce them to raise freight rates. Thus, the co-movement of tariffs and freight rates identifies market power.

[27] Interestingly, when regressing transport prices on road condition, they found negative and significant effects in East Africa, but insignificant or positive effects in West and Central Africa (Table 4.3 p. 42), suggesting pricing formulas based on anticompetitive arrangements rather than marginal costs in those regions.

[28] They collected data on costs (Vehicle Operating Costs (VOCs), transport costs incurred by transport providers) and prices paid by end users from a sample of trucking companies operating across the continent. They then simulated the effects of a reduction in: (i) fuel price; (ii) informal payments; (iii) reduction in border crossing time; (iv) rehabilitation of corridors. Their simulations showed that for West Africa (and to a lesser extent Central Africa), a reduction in fuel prices and a rehabilitation of roads would have no effect on prices paid by end users because of barriers to entry. By contrast in Eastern Africa, the same policies would reduce prices paid by end users.

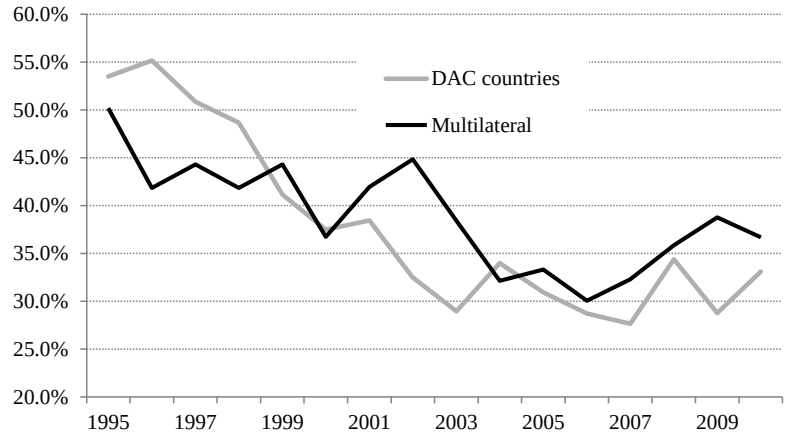
improvements have a bigger impact on export growth than tariff reductions. Helbe et al (2009) find that exporter and importer transparency is more important than trade barriers for bilateral export of differentiated products. Francois and Machin (2007) find that those infrastructures are more important than trade barriers for bilateral export growth. Portugal and Wilson (2008) find that a substantial but feasible improvement in trade facilitation would be equivalent to a tariff reduction of 8 percent in Africa. Wilson et al (2003) performed the same exercise in APEC and they found that the required improvements in trade facilitation indicators are relatively small compared to the equivalent tariff reduction. Hoekman and Nicita (2008) conclude that the impact of reducing transaction costs at and behind the border will have a greater payoff than further reduction in tariff and NTMs. Since these reductions do not require multilateral negotiations and can be done unilaterally, there is great scope to enhance growth opportunities for developing countries "while Doha sleeps". In the same vein, Djankov et al (2010) find that in LDCs reducing delays by ten days would give more export growth than liberalization with EU or USA.

What do we learn about prioritization? Most studies above have tried to include different dimensions of trade facilitation to observe which ones would be the most constraining, but no common pattern emerges. Wilson et al 2003 found that individual APEC members differ in terms of which trade facilitation reform would be more profitable. Portugal and Wilson (2008) find a slightly more important impact for reforms improving indicators captured in the LPI index than for reforms reducing trade costs captured in Doing Business, though overall both dimensions appear important. Francois and Machin (2007) find that their infrastructure index (generated with a PCA on 8 infrastructure indicators) is strongly more important for bilateral export growth than their institution index (also generated with a PCA on 7 indicators from Fraser institute). On the contrary Balchin and Edwards (2008) on African countries, find no effect of their infrastructure index (generated with PCA) on export propensity (dummy 0/1) while their index for legal environment and micro level supply constraints are strongly significant. Hoekman and Nicita (2008) find that policies affecting what is captured in LPI have more effect on bilateral export than those affecting what is captured in DB.

4. Does AFT Make a Difference? Inspecting the Aggregate Data

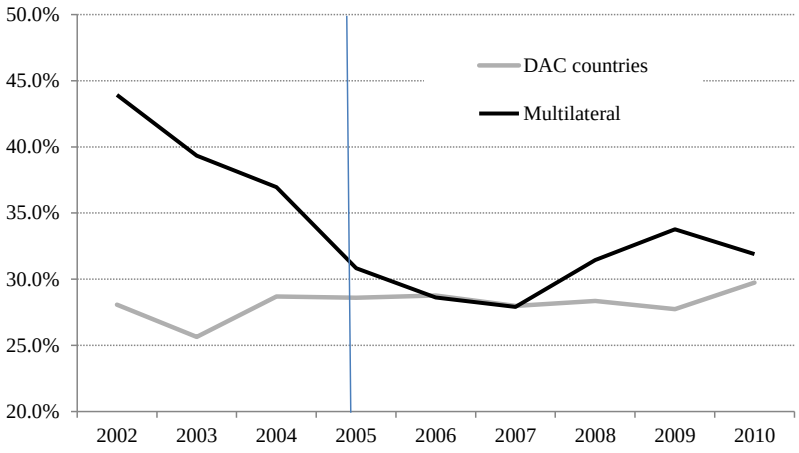
As reflected in the individual Aid-for-Trade at a Glance country sheets (OECD 2000) and further noted by Cadot and Newfarmer (2012), trade negotiators, in particular from developing countries, have had a tendency to assess success in AFT by the amounts. As shown in Figure 1, by the commitment measure, 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). Looking at disbursement trends in Figure 2 reveals that the increase comes later reflecting disbursement delays Observe that the share of bilateral Aid from DAC countries to the AFT agenda did not increase indicating that the "big push" came mostly from multilateral agencies.

Figure 1
Aid for Trade Commitments, 1995-2010 (share in Total ODA)



Source: Authors' calculations using OECD CRS database

Figure 2
Aid for Trade Disbursements, 2002-2010 (share in Total ODA)



Source: Authors' calculations using OECD CRS database

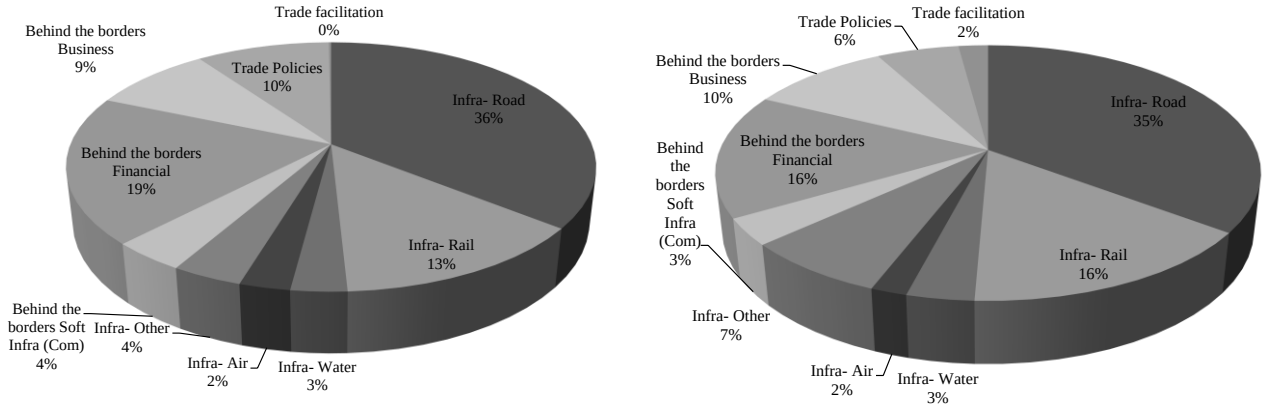
4.1. How Much? To Whom?

However, commitment or disbursement statistics are plagued with measurement issues making it difficult to evaluate their impact. As discussed in the annex, the Creditor Reporting System (CRS) does not provide information about trade-related technical assistance and trade development which was previously collected in the joint OECD-WTO Trade Capacity Building Database (TCBD). Modifications to the CRS classification to match more closely our operational definition of AFT in figure 1 are described in annex tables A2 and A3.

Figure 3 shows the evolution between 2002 and 2010 for the main components or our classification. It shows little change in the allocation of AFT over the period (this would also hold even if we kept the original classification).

Figure 3
Components of Aid for Trade, 2002-2010

(a) 2002 (b) 2010



Source: Authors' calculations using OECD CRS database

Since the controversial paper by Burnside and Dollar (2000) arguing that aid had a positive impact on growth only in countries that followed 'good' policies (as measured by the World Bank's CPIA measure), the effects of aid on growth have been widely studied. So has aid allocation in a voluminous literature aimed, inter alia, at finding whether "deserving" countries in terms of governance and capabilities got more aid than others; results were, by and large, ambiguous.[29]

The allocation of AFT has been much less studied, and results are not only few but plagued by identification issues. Samy and Imai (2010) regressed AFT disbursements (also from the CRS) on country characteristics including GDP per capita, infant mortality, population, as well as democracy, openness and infrastructure indices in a panel of countries over 1997-2006. Fixed-effect estimates returned somewhat unintuitive results: country characteristics that seemed most robustly associated with AFT allocations (infant mortality, democracy, GDP per capita) had only indirect linkages with trade gaps, whereas infrastructure indicators such as roads or energy supply were insignificant, and more concentrated exports correlated with lower inflows. These hard-to-interpret results might have to do with identification issues, with time-invariant country fixed effects not sufficient to control for omitted variables.[30] Gamberoni and Newfarmer (2009) constructed a

[29] For instance, Alesina and Dollar (2000) found that political rights had a significant effect on aid receipts. Alesina and Weder (2002) found no evidence that corruption reduced aid inflows, but Dollar and Levin (2004) found a positive effect for governance.

[30] For instance, a period of trouble might result both in more aid flows aimed at rebuilding infrastructure and institutions and more concentrated exports, as the mining sector, which operates in well-guarded enclaves, is typically more robust to instability than light manufacturing. Also, estimation did not control for selection or censoring of AFT inflows at zero.

composite measure of "AFT demand" at the country level based on proxies for under-trading and governance. Regressing AFT inflows from the OECD's CRS database on this indicator on a cross-section of countries, they found a positive association, although with wide dispersion. Tadasse and Fayissa (2009) regressed U.S. AFT outflows on various country characteristics. Notwithstanding the usual identification issues, they found that more AFT tended to go to close-by countries with strong trade ties with the U.S.[31]

4.2 Any effect?

As a first pass, we ask whether there is any correlation between export growth and lagged AFT commitments.[32] Clemens, Radelet, and Bhavnani (2004) divide aid flows into three categories: Short term emergency aid likely to be negatively correlated with growth; aid for health, education the environment and support for democracy that affects growth over a long period of time; and aid that plausibly could stimulate growth in the short term, including budget and balance of payments support, investments in infrastructure, and aid for productive sectors such as agriculture and industry. AFT falls mostly under this last category so we would expect to see some correlation between average disbursements over a five-year period and average export growth—or some proxy for trade-costs reduction—over the next five-year window.

Figure 4 provides a very simple check on whether such a correlation is visible to the naked eye. We split the set of AFT recipients by the median into two cohorts, "low recipients" and "high recipients", based on average 2000-2005 receipts. We would want to see higher export growth in the latter group than in the former over the next five-year window (2005-2010), the lag being to leave room for delayed effects. In order to get some more information out of the data, Figure 4 looks separately at each quintile of the (baseline) export/capita distribution. 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 do we see a positive export-growth differential between high- and low-recipients (Panel a). On the possibility (see below) that AFT will have an indirect effect on export performance by working primarily through improved logistics markets, Panel b carries out the same exercise for the time to export, with similarly disappointing results.

So much for AFT fostering convergence. Of course, our exercise is rough and ready, and many confounding influences and channels of reverse causality should be filtered out before any firm conclusion is reached, preferably using some econometrics. We now turn to some of the literature's results.

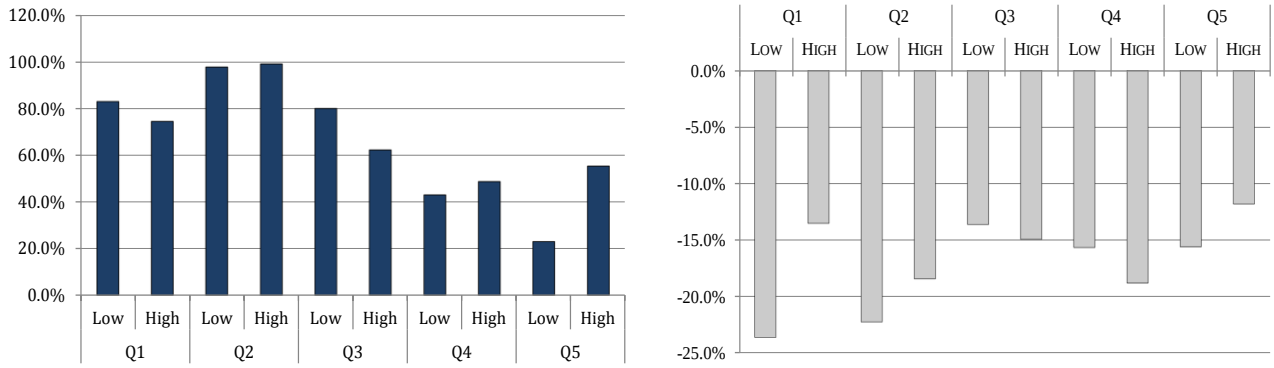
[31] These results echo those of Alesina and Dollar (2000) who find that disbursements for European countries with former colonies go mostly towards the former colonies. Berger et al. (2013) show that during the cold war, CIA interventions led to a surge of imports towards the US in industries in which the US had a comparative disadvantage.

[32] An early strand of the literature explored whether AFT positively affected the *donor*'s exports (up to the early 1990s, over half of all bilateral aid was at least partially tied to donor exports). Using a gravity equation, Wagner (2003) and Nelson and Juhasz Silva (2008) found a positive association, although the effect was small; Osei, Morrissey, and Lloyd (2004), using a gravity equation in first differences for a panel of four European donors and 26 African recipients, found an unstable and insignificant impact of aid on exports from donor to recipient.

Figure 4

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: Authors' calculations using OECD CRS database and WDI

Using the OECD's CRS,[33] as suggested by (4), Cali and te Velde (2011) regress trading costs and aggregate export value on lagged AFT disbursements and control variables, on a panel of developing countries. Identification is based on recipient fixed effects (FE) and instrumenting AFT flows with the Freedom House's index of civil liberties, the authors arguing that the Millenium Challenge Corporation explicitly uses that index as an input in their aid allocation mechanism. For aid to infrastructure, coefficients are significant in some specifications, but with very limited robustness. As for aid to productive capacity, it fails to correlate with exports whatever the specification, estimator, or lag structure. As for results on sectorally-targeted aid, they tend to confirm the profession's longstanding skepticism about targeted support. Cali and te Velde (2011) find significant effects only in some specifications, and they vanish as soon as comparative advantage is controlled by country-sector fixed effects.

Brenton and von Uexkul (2011) find that, in a simple before-after comparison, sectors that receive aid support perform better, but a difference-in-differences regression of country-sector exports on aid flows controlling for heterogeneity through matching does not show significant effects (in particular once outliers are eliminated), suggesting that crude comparisons that fail to control for aid endogeneity pick up reverse causation. The only significant results obtained on sectorally-targeted AFT concern aid to service sectors. Ferro, Portugal and Wilson (2011) combine aid to services with input-output data to identify the effect of aid on downstream manufacturing sectors, reasoning that reverse causality (from sectoral performance to aid flows) would be eliminated by looking at aid to upstream services. Using a country-sector panel for 1990-2008, they find significant effects for aid to banking services and energy but, interestingly, no effect for aid to ICT or transport services (distinct from transport *infrastructure*).

[33] Trading costs are measured by the trading across borders indicators of the Doing Business database.

Building on empirical results discussed earlier in this survey that suggest a robust correlation between infrastructure quality and export performance, Vijil and Wagner (2012) look for the effect of infrastructure-aid commitments on an index of infrastructure quality composed of roads and telecom densities in a cross-section of 91 countries for which they take average values of all variables over 2002-2007. They control for overall ODA inflows, geography and institutions (proxied by a rule-of-law index), and deal with endogeneity and measurement error by instrumenting aid to infrastructure by the number of privatizations in the infrastructure sector between 2000 and 2007. They find that when all country controls are included , the quality of infrastructure is significantly positively correlated with aid to infrastructure in all 2SLS specifications.

The literature thus far used recipient or donor/recipient characteristics only as controls, leaving aside the important question of synergies between aid flows and other forms of development assistance. Vijil (2012) explores the complementarity between AFT and regional integration, using Baier and Bergstrand's measure of trade integration (Baier and Bergstrand 2009). She uses a panel gravity on 178 countries over 1995-2005 where she introduces AFT inflows for the exporter and importer linearly and interacted with the bilateral integration term. She finds that bilateral trade correlates significantly with AFT levels in the exporter and importer countries, but also that the effect is reinforced by integration, in particular on the importer side. The effect is strongly significant for South-South and North-South trade, possibly reflecting complementarities between aid flows and regional cooperation on NTBs and transit, or the policy stability that comes with regional integration.

5. Impact evaluation in trade

Evaluation is part of a broader results-assessment nexus comprising monitoring, a continuous process tracking whether program implementation follows established rules and procedures, and evaluation, a periodic exercise addressing performance issues, including causal links between program and outcomes and possibly cost issues as well. The focus on causal links ("attribution") is what distinguishes impact evaluation (IE) *stricto sensu* from other forms of evaluation. The typical questions that one would want to address through impact evaluation are:

1. Are there changes in the observed outcomes of beneficiaries of program *P* that can be attributed to the program? Are those changes replicable?
2. What components of program *P*'s treatment ("treatment arms" in the IE jargon) are most effective?
3. How do program *P*'s benefits compare with its costs? Is it the most cost-effective among alternatives?

Of course, a single impact evaluation is unlikely to provide final answers to all three questions. In what follows, we will try to provide a realistic overview of what can be expected from IE, in particular in a trade context, and what cannot.

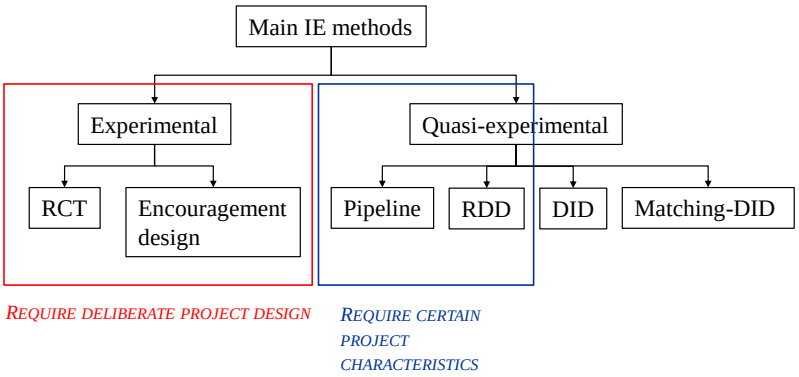
5.1 Methods and interpretation: A brief overview

5.1.1 Data and methods

Although impact evaluation methods are familiar to most economists, in particular labor economists who have used them for over four decades, a brief overview is in point to motivate our later discussion of how those methods can address the needs of AFT evaluation. Notwithstanding recent controversies, the essence of impact evaluation is the use of a control group to provide a counterfactual to treatment effects; it is not randomization, which is one method among several to ensure comparability of the treatment and control groups.

Figure 6 describes the two families of methods used in IE: experimental design (randomization) and quasi-experimental design (econometrics).[34] They are represented, together with their main members, in Figure 5. Starting from the left, randomized-control trials (RCTs), the "gold standard" of IE, rely on randomized assignment (i.e. on the law of large numbers) to ensure comparability of the treatment and control groups. In principle, randomized assignment ensures that potential individual outcomes do not correlate with treatment status. Put differently, provided that sample size is sufficient to ensure that the treatment and control groups are similar in their pre-treatment attributes, randomization ensures *internal validity*, i.e. the ability to establish a causal relationship between outcomes and treatment. By contrast, *external validity*, i.e. the ability to draw inferences that are valid out of sample, requires the overall sample (of treated and control individuals) to be representative of the wider population of "treatable" individuals. This is usually difficult to establish and is a key drawback of IE, to which we will return.

Figure 5
Impact evaluation methods



[34] For a rigorous overview, see Duflo, Glennerster and Kremer (2006). See also the very pedagogical approach in Gertler et al. (2010) or Khandker et al. (2009).

In practice, randomization is rarely part of the culture of government agencies and can easily be perceived as ethically difficult or unrealistic.[35] However, when programs are oversubscribed or during their pilot phases, randomized assignment should pose no particular ethical problem as it is a fair way of allocating scarce resources. Also, micro-programs to encourage women entrepreneurs in export-oriented sectors, sometimes run by NGOs, could be evaluated like other anti-poverty programs through RCTs.

When a program has universal eligibility or no control over who enrolls, randomized assignment is not possible; randomized *promotion* (so-called "encouragement design") is then an alternative. Randomized promotion combines differences in outcomes with differences in take-up rates between the treatment and control groups (due to the fact that the program was not promoted to the control group) to infer treatment effects.[36] Under encouragement design, the power of the treatment-effect test depends not just on sample size but also on the promotion's effectiveness, which must generate substantial differences in enrolment rates between the two groups. For instance, encouragement design can be used to evaluate the use of inputs (fertilizers, pesticides etc.) in export-oriented agriculture.

Both methods require deliberate program design from the outset that goes beyond just collecting baseline data. However, in both cases (micro-entrepreneurship programs or agricultural-support ones) there is an issue about the value added of impact evaluation, especially given its cost (an issue to which we will return later on). Micro-entrepreneurship programs are typically relatively low-scale and low-stakes, and a lot of ink has already been spilled on whether they make a difference or not (see e.g. Banerjee and Duflo 2011). Would the n^{th} evaluation of a micro-program make a difference in the policy debate? As for the effect of using fertilizers, agronomists on the ground have pretty good knowledge of what works and what doesn't. What radically new information would an IE uncover?

Programs that have not or could not be designed for IE, which may well be those we are most interested in, can nevertheless be evaluated, under certain conditions and depending on their characteristics, using various quasi-experimental (QE) methods.

For programs with progressive phase-in, pipeline methods use eligible but not-yet-enrolled individuals as controls. This has the advantage of alleviating some of the usual selection problems, as unobservable characteristics correlated with the willingness to enroll can be expected to be the same for enrolled individuals and those in the pipeline.

For programs with a well-defined cutoff in terms of a certain observable attribute (say, SME support programs where firms must not have more than a given annual turnover), regression-discontinuity design (RDD) uses outcome differences around the cutoff point for the identification of treatment effects. That is, in the SME example, ineligible firms immediately above the cutoff,

[35] There are prominent counter-examples. For instance, Mexico's celebrated PROGRESA anti-poverty program (deployed in May 1998) was randomly assigned for rigorous evaluation from the outset. As a result, it was widely studied and has set the standard for such programs worldwide. The first evaluation was carried out by IFPRI at the request of the Government of Mexico in 1999. See http://www.ifpri.org/dataset/mexico-evaluation-progresa.

[36] In econometrics terms, randomized promotion serves as an instrumental variable for enrolment.

which are almost identical to eligible firms immediately below, are used as the counterfactual. The sample-size requirement is crucial here, as inference is based on a small « bandwidth » around the cutoff. Moreover, the external validity problem is particularly severe as firms far away from the cutoff have no natural counterfactual.

Finally, programs that have not been designed for evaluation and have neither pipeline nor cutoff can be evaluated using difference-in-differences (DID) regression, with or without matching. DID methods compare the before-after variation in outcomes (first differences, usually of log-levels, which means growth rates) for the treatment group with the before-after variation for the control group. DID regression filters out two types of confounding influences: common trends and (some) heterogeneity between individuals.

As for common trends, suppose that all firms experienced reduced exports because of a macroeconomic downturn at the time an export-promotion program was deployed. A crude before-after comparison of the treated firms' exports would suggest that the program *reduced* them. The confounding influence of the downward trend is filtered out by the comparison of export growth for treated vs. control firms (the second difference in the DID), provided that both are subject to the same trend.

As for heterogeneity, suppose that the treatment group has larger firms that export more, making the two groups not directly comparable. First differencing filters out this confounding influence, which is on outcome *levels*. Suppose further that the larger firms in the treatment group tend to grow less because of convergence. This is a confounding influence on outcome *growth* that is not filtered out by first differencing. It is (or can be) filtered out by controlling for individual covariates in the regression (here initial size), provided that the data set is rich enough.

A further problem may arise because of individual heterogeneity in the degree of responsiveness to the treatment. This one can be reduced (if not eliminated) by so-called "propensity-score matching", which consists of constructing a control group whose probability of treatment, based on observable attributes, is as close as possible to that of the treatment group.[37] Suppose that it so happens that firms with a large proportion of skilled labor are both more likely to sign up for a technical-assistance program and more responsive to it. DID estimates will then suffer from endogeneity bias, as treatment status, a RHS variable, will be correlated with the error term. Matching-DID estimates can correct for this, provided that the individual attribute correlating with both enrolment and responsiveness is both observable and time-invariant. If it was related to, say, changes in firm management, matching would not help.

DID and matching-DID methods can be further refined by testing whether "placebo" treatments have any effect (say, by introducing faked treatment dates).

To sum up, because of their common reliance on a control group and variation in outcomes before and after the treatment, practically all of these methods require two basic ingredients:

[37] Practically, what that means is that the econometrician regresses treatment status (a zero-one variable) on individual characteristics for the whole sample (treated and nontreated individuals), retrieves predicted treatment probabilites, and then pairs each treated individual with those in the control group that have the closest predicted treatment probability (one or many, depending on the matching algorithm).

- A baseline survey covering sufficiently many individuals and sufficiently many attributes of those individuals
- A group of individuals left untreated, whatever the assignment method.

Aside from scholastic debates about whether randomization is the alpha and omega of evaluation or not, if all projects routinely had those two features, assessing their effects would be a lot easier.

5.1.2 Interpreting IE results

Even with adequate data and basic design, conceptual and statistical issues will always plague impact evaluation. One of them is externalities. What if the program's treatment spills over to the control group? For instance, export promotion could lead firms to make easily imitable product or country expansion decisions, in which case the control group's outcomes would quickly converge to those of the treatment group, leading the evaluator to believe that there was no lasting treatment effect.

The problem is more than just one of identification. In areas like health, education, gender or poverty eradication, policy is aimed at the needs of vulnerable individuals—women, children, individuals with poor health—and the government may care about their personal well-being even in the absence of spillovers. By contrast, in trade, we are (usually) dealing with firms. Whereas improving a child's wellbeing is a valid policy objective in itself, improving the welfare of a company's shareholders is not. Thus, in a trade-related environment, a government program can be cost-effective (the third of our basic IE questions) in an economic sense only if it addresses a well-defined market failure. If it were simply to channel the taxpayer's money to privately-appropriated benefits, it could not be cost-effective, as the distortions created by taxation would not be offset by any social benefits. But, by construction, treatment effects pick up only (or essentially) privately appropriated benefits, since social benefits would affect the control group as well.

Thus, IE should be used cautiously as a policy-evaluation tool, and preferably in conjunction with some direct measure of spillovers. Consider again our export-promotion program. If the IE failed to show any significant treatment effect, it could be that (i) the program was ineffective, (ii) the program was effective but its benefits fully spread to the control group, or (iii) the test did not have enough power given the sample's characteristics. In case (i), the program should be phased out; in case (ii), it should not, *provided that the benefits were valuable to society at large or could be taxed back*; in case (iii), we wouldn't know. Thus, the "no treatment effect" result is fundamentally ambiguous. If, by contrast, the IE did return a significant treatment effect, the null of treatment ineffectiveness could be rejected, but there would still be no justification for using the taxpayer's money for it. Instead, it should be implemented on a full-cost recovery basis.

To see this, suppose now that the market failure was credit rationing, exporters being denied access to trade finance because of asymmetric information, moral hazard, the absence of institutional arrangements for collateral, or any other market failure. The presence of a market failure would provide a justification for government intervention, say, in the form of a guarantee fund. In that case, the benefits would be fully appropriable by exporters, which would return significant

treatment effects in an IE. But again, the program should be implemented, if possible, on a full-cost recovery basis.[38]

5.2 What IE for what intervention?

5.2.1 Data

The previous section's discussion of available IE methods clearly suggests that "clinical" policy interventions are more amenable to IE design than universal policy reforms. Thus, traditional trade-policy reforms such as reductions in tariff or non-tariff barriers are, by construction, outside of IE's scope and should rather be evaluated using traditional econometric methods, provided that usual identification problems such as the endogeneity of policy reforms can be properly addressed by instrumentation or through a natural experiment.

A good example of this approach is Estevadeordal and Taylor (2009), who used countries that liberalized tariffs on capital and intermediate goods after the Uruguay Round as a treatment group and countries that did not as controls in a DID regression assessing the effect of trade liberalization on growth. In order to deal with the treatment's potential endogeneity, the identification strategy could not rely on the usual trick of using "institution-quality" proxies, since those are largely time-invariant and thus unsuitable to a DID framework. Instead, it relied on a historical argument according to which countries that underwent a less severe GDP drop during the Great Depression of the 1930s would be more inclined to liberalize later on. While intuitive, this instrumentation strategy clearly relies on a long causal chain that is fraught with uncertainties and confounding influences. It is those influences that narrower IEs try to escape.

Turning to narrower policies with better scope for applying IE techniques, the twin requirements of adequate project design and a "clinical" (or "targeted") nature of the intervention suggest the typology of projects and evaluation methods shown in Table 1. Targeted programs such as technical assistance, export promotion, etc. could be amenable to RCT design provided that the decision to randomize assignment was taken ex ante (upper left-hand cell in Table 1). In practice, only a minority of them can be expected to be RCTs. If not (upper right-hand cell in Table 1), QE design is appropriate. In that case, ideal data requirements will typically include

- *Trade data* at the transaction level from customs, which is available from ASYCUDA raw files. The data can be easily anonymized by deleting firm names and keeping only TINs (tax identification numbers) and will provide information on firm-level outcomes;
- *Program data* including enrolment status, dropouts, and rejects;
- *Firm characteristics data* from an industrial survey (typically balance-sheet information including employment, turnover, age, as well as ownership, number of establishments, etc.); of course, the survey's key for the classification of firms should be compatible with that of customs data for reconciliation, which precludes the use of "dummy" firm identifiers.

[38] In small developing countries, a guarantee fund may also have systemic effects on the stability of the country's banking system, which is an externality. In practice, some guarantee funds deployed by donors such as Agence Française de Développement are intended to be self-financing, but the subsidy component appears difficult to phase out.

Clearly, these data requirements are heavy and raise confidentiality issues; whether the data will actually be made available to the evaluation team by government authorities depends on interest (buy-in) for the IE's results, donor involvement, and quality of the dialogue.[39]

Table 1
Boundaries of impact evaluation

	Evaluation built into program design	Evaluation not built into program design
Targeted (typically trade competitiveness-related e.g., matching grants for producers for technology upgrading or export business plans; export credit guarantees for producers)	RCT is feasible; Quasi-experimental methods are a possible alternative	RCT is infeasible; Quasi-experimental methods are feasible
Non-targeted (typically trade facilitation-related: e.g., customs reform, port improvements; but also some trade competitiveness -related: support for producer organizations or other institutional reforms)	RCT is typically infeasible; Quasi-experimental methods are more appropriate; Some methods of targeting can be introduced (phase-in, staggered implementation)	All IE methods are difficult; before-after comparisons may be only alternative

Notes: RCT: Randomized control trial; QE: Quasi-experimental. QE methods are matching, difference-in-differences, instrumental variables, or regression discontinuity design (see figure 6).
Source: Cadot, Fernandes, Gourdon, Mattoo (2011, Table 2)

Non-targeted programs such as customs reform (lower cells in Table 1) will typically be more difficult to evaluate with methods other than crude before-after comparisons, although progressive phase-in will allow some scope. For instance, suppose that border posts are modernized one at a time, with improved facilities, performance contracts for customs officers, and various measures to improve their productivity (see Cantens and Raballand 2011 for an example in Cameroon). One immediately thinks of an analogy with pipeline methods where border posts to be modernized are used as controls for those being modernized. However, in many low-income countries, there are few border posts and they are typically different from one another in terms of volume, type of traded commodities, type of firms, etc. For instance, there will be a border post linked to a port and smaller ones for overland trade. This would make border posts bad counterfactuals for one another. However, if one shifts the analysis to the level of the transaction, a transaction using one border

[39] The World Bank has recently launched the « Exporter Dynamics » project which aims at collecting precisely this type of data (at least the customs data) from Customs administrations around the world. However, sharing the data with researchers has proved a difficult and laborious process because of the confidentiality issues involved.

post may be a good counterfactual for another one in a similar commodity, and performed by a similar firm, using another post, even if those posts are not comparable on aggregate. If, say, transactions in a reformed border post have undergone a reduction in clearance delays around the time of reform while similar transactions in other border posts have not, one may be able to identify a credible treatment effect.

Thus, the availability of transaction-level data can open up possibilities for evaluation that would not be there without them, and this even in the bottom-right-hand box of Table 1 where no deliberate effort was made ex ante to make project design amenable to evaluation.

5.2.2 Implementation

In practice, efforts to generalize the use of IE in trade interventions face two types of constraints: incentives and resources.

In terms of incentives, an IE risks slowing down project roll-out and diverting managerial attention for results that are unlikely to be available within a manager's tenure horizon; and if they did, they might do more harm than good. The way typical incentive structures work, the accumulation of smoothly-run projects satisfying internal success criteria (full and timely disbursement, completion of planned tasks, and some evidence, say from focus groups, that beneficiaries are happy) is the stuff of successful careers. By contrast, a single project attracting controversy may do permanent damage. An IE saying that a project had an impact will just confirm what the project manager was probably claiming anyway; one that would say the project had no impact would dangerously contradict his claims and trigger controversy.

In order to be incentive-compatible, IE should be used only to generate new knowledge and should be fully decoupled from the evaluation of project managers. However, it is not clear that an organization could make such a claim credible, as it would obviously suffer from a time-inconsistency problem. Thus, the decision to carry out an IE or not should not be taken by project managers. It should be part of the project's design upstream of the project manager. But even if that happened, IE might affect the manager's decisions ex post by discouraging experimentation. If the project manager could choose between two "treatment arms", a proven but un-innovative one and a new but risky one, the perspective of an IE could lead him/her to avoid the innovative/risky one.

As for funding, the basic issue is that IEs have tended to be on the expensive side, although there may be scope to reduce costs. This is illustrated by
Table 2, reproduced from Gertler et al. (2011). In a sample of IEs of World-Bank-supported projects over recent years, the average IE cost is over $700'000. In an interview with the authors, the head of the World Bank's DIME unit reckoned that with less than $300'000, it would be difficult to think of a serious IE. What is so expensive?
Table 2 shows that most of the cost is accounted for by data collection, which is very expensive when it comes to household surveys. For instance, in an interview with the authors, specialists in the World Bank's LSMS division estimated the average cost of a household survey at about $300

per household, including repeated visits and all related costs. Still, even at that rate, the particularly expensive IE in Malawi ($1.8 million, of which $1.3 million for data collection) could cover 2'000 households in a two-period panel, a large sample for an IE. Combining this with travel costs of about $83'000 ($270'000 for the Benin IE in Table 2) may lead the reader to think that there is scope for reducing IE costs to more palatable figures.

Table 2
IE costs, selected World Bank supported projects

	Country	Total cost	Travel	World Bank staff	Consultants (national and int.)	Data collection (inc. Field staff)	Dissemin.
Poverty Reduction Support Credits and Maternal Health Performance	Benin	1690,000	270,000	200,000	320,000	840,000	60,000
Pay Reform for School Teachers	Brazil	513,000	78,000	55,000	105,000	240,000	35,000
Nadie es Perfecto Program to Improve Parenting Skills	Chile	313,000	11,500	—	35,500	260,000	6,000
Paying for Performance in China's Health Sector: Evaluation of Health XI	China	308,900	60,000	35,000	61,000	152,900	—
National Rural Employment Guarantee Program	India	390,000	41,500	50,000	13,500	270,000	15,000
School Health and Nutrition: the Role of Malaria Control in Improving Educat	Kenya	652,087	69,550	60,000	103,180	354,000	65,357
HIV Prevention Campaign for the Youth: Abstinence, Fidelity and Safe Sex	Lesotho	630,300	74,300	9,600	98,400	440,000	8,000
CCT, Schooling, and HIV Risk	Malawi	1842,841	83,077	144,000	256,344	1359,420	—
Contigo Vamos por Mas Oportunidades Program in the State of Guanajuato	Mexico	132,199	2,660	50,409	—	80,640	1,150
Randomized CCT Pilot in Rural Primary Education	Morocco	674,367	39,907	66,000	142,460	426,000	—
Learning and Growing in the Shadow of HIV/AIDS: Randomized ECD Program	Mozambique	838,650	86,400	31,000	62,500	638,750	20,000
Training of Community Distributors in the Prevention and Treatment of Malar	Nigeria	1024,040	64,000	35,000	106,900	817,740	—
School Health and Nutrition: the Role of Malaria Control in Improving Educat	Senegal	644,047	61,800	60,000	102,890	354,000	65,357
CCTs to Prevent HIV and Other Sexually Transmitted Infections Tanzania	Tanzania	771,610	60,000	62,000	100,000	518,611	30,999
Average		*744,646*	*71,621*	*66,031*	*115,975*	*482,290*	*30,686*

Note: CCT: Conditional Cash Transfer; ECD: Early Childhood Development
Source: Gertler et al. (2011), Table 10.3

If IEs are so expensive, how did they spread in so many areas of development aid? The answer is in Table 3. In areas such as social development, the amounts in Table 2 may not seem outrageous compared to the sheer size of the programs. Gertler et al. (2011) show that, for a sample of World Bank-supported programs for which IE and program cost data was available, IE costs represented on average between 4% and 5% of total program costs, ranging between 0.2% and 13.3% (Table 3). This is because project costs in the sample ranged between $11 million (Rwanda) and $86 million (Colombia).

Trade-related projects rarely attain such levels. If we take DIME's estimate of a minimum of $300'000 for a feasible IE, a ratio of 5% would require a project of $6 million. By the standards of World Bank projects in social development, poverty, or health, this is a small project. By the standards of trade-related assistance, this is very large.

Table 3
IE costs as a fraction of program cost, selected World-Bank supported programs

		Program cost	IE cost	Percent
Social Safety Net Technical Assistance	Pakistan	60'000'000	2'000'000	3.33
Social Fund for Development 3	Yemen, Rep.	15'000'000	2'000'000	13.33
Social Protection Project	Panama	24'000'000	1'000'000	4.17
1st Community Living Standards	Rwanda	11'000'000	1'000'000	9.09
Social Protection	Jamaica	40'000'000	800'000	2.00
Social Sectors Investment Program	Dominican Rep.	19'400'000	600'000	3.09
Migrant skills Development and Employment	China	50'000'000	220'000	0.44
Social Safety Net Project	Colombia	86'400'000	130'000	0.15
Average				*4.45*

Source: Gertler et al. (2011), Table 10.2

In order to provide a rough order of magnitude of the size of AFT projects, Table 4 displays mean and median annual commitment size in 2010, by country, from the OECD's CRS database, cumulated over all donors (23 DAC countries and 10 multilateral). This can only provide a very rough order of magnitude since (i) several donors will typically be active in a given country/area, (ii) a donor's annual commitment to an area may cover several projects, and (iii) individual projects can extend over more than a single year. Be that as it may, median commitment size for trade policy and regulations, by country, was in 2010 $700'000. With a conservative estimate of 3 donors per country and one project per donor, we would be slightly above $200'000 per project, i.e. below the "minimum-cost" IE at current rates. For banking and financial services, it was $2.9 million, or $1 million per donor under the same assumptions. Clearly, these amounts are orders of magnitude below the amounts typically committed to poverty reduction, social development, or health, and would make it difficult to spend on IEs the kind of amounts shown in Table 2.

These very rough calculations help explain the slow spread of IE in trade-related assistance and suggest that IE templates must be adapted to the area of trade assistance in order to make IE an acceptable proposition for donors. Clearly, QE methods using statistical data instead of original household surveys are the way to go. We now turn to a few examples of recent IEs in that spirit and how they have contributed to our understanding of the effectiveness of trade interventions.

Table 4
Annual commitment amount, 2010, by country and sector (million dollars)

	Mean	Median
Transport & storage	110.3	28.3
Energy	84.7	16.2
Agriculture	49.1	17.0
Banking & FS	15.3	2.9
Business services	11.9	4.1
Industry	11.7	3.6
Forestry	7.3	0.8
Mineral resources	6.9	0.4
Trade policy & regulations	3.7	0.7
Fishing	2.6	0.4
Communication	2.6	0.6
Tourism	1.4	0.2

Source: Authors calculations from OECD, CRS

5.3 Early results: Does export promotion make a difference?

So far, there have been very few impact evaluations of trade-related interventions, and only a thin, "early" literature can be reported on. However, the performance of export-promotion agencies, which is one of the few areas of "clinical-type" interventions that have been extensively studied, provides a good testing ground to evaluate the contribution that IE—defined the way we have defined it in this survey—can bring to policy debates and dialogue with developing countries.

Export-promotion agencies (EPAs) aim at overcoming informational barriers faced by existing and potential exporters by helping them with market prospection and promotion, and sometimes through technical assistance as well (product and firm development). In the early 1990s, on the basis of broad cross-country experience, a World Bank study (Keesing and Singer, 1991) concluded that in most developing countries, EPAs had failed to make any difference in terms of export performance. Factors that contributed to their inefficacy included under-funding, government meddling in the management, lack of private-sector involvement, and most important of all, overall policy environments that, until the wave of reforms in the 1990s, were strongly biased against exports.

By the mid-2000s, the policy environment in many developing countries had changed drastically, and some of the changes suggested by Keesing and Singer had been implemented. A new, cross-country study by Lederman, Olarreaga and Payton (2010) using a survey of EPAs, suggested that, contrary to perceptions, the return to additional funding was in fact very large. The study's identification strategy illustrates the difficulties facing cross-country econometrics when it comes to impact evaluation. Clearly, the *presence* of an EPA is endogenous to trade performance (a selection problem) and so is the *intensity* of promotion, i.e. the EPA's budget. The authors use Heckman's estimator to deal with selection, using GDP per capita and aid per capita as excluded variables in the selection equation (although the case for excluding per-capita aid from the export equation could be debated). As for endogeneity of the EPA's budget, they use the agency's age and time to election as excluded instruments. With typical EPA budgets standing at about a thousandth of export value and elasticities of exports to EPA budgets between 0.05 and 0.09, raising an EPA's budget by 1% from a baseline $1 million (a $10'000 increase) would raise exports by $500'000-$900'000 (0.05%-0.09% of $1 billion), or a return of $50 to $90 to the dollar. Such a too-good-to-be-true rate of return suggests that the cross-country estimation may pick up confounding influences, which is not overly surprising since the nature of the data prevented the authors from using panel-data econometrics (the EPA survey was carried out for a single year).

Partly in reaction to the traditional identification difficulties facing cross-country studies, partly as a result of the increasing availability of firm-level data, a new literature surveyed in Volpe (2011) has turned to "clinical" (firm-level) evaluation of EPAs. For instance, using DID estimation at the firm level, Alvarez and Crespi (2000) evaluated Chile's EPA, PROCHILE, and found that its activities had an impact on the beneficiaries' number of export destinations, although not on their number of export products. Since then, a number of firm-level studies have shown that export promotion seems to be more successful at affecting the performance of established exporters than at encouraging non-exporting firms to start exporting (Bernard and Jensen, 2004; Görg, Henry and Strobl, 2008; Girma, Gong, Görg and Yu, 2009), in accordance with the literature on heterogeneous firms and trade, which suggests that exporters differ from non-exporters in terms of productivity and a host of other firm characteristics (see, e.g. Bernard, Jensen, Redding and Schott, 2007), which export promotion activities may not be able to offset. Moreover, the impact seems stronger along the extensive margin than along the intensive one (Alvarez and Crespi, 2000, Volpe and Carballo, 2008, Cadot, Fernandes, Gourdon and Mattoo 2012). Thus, assistance may be more successful in helping firms overcome hurdles to break into new markets (product- or destination-wise) than in ramping up export volumes.

Did this literature produce any insight that the cross-country literature did not? On one hand, it did not *overturn* the qualitative result of Lederman, Olarreaga and Payton (2010). They had found that EPAs do make a difference; this finding is upheld. On the other hand, the result is qualified and refined in a number of ways. For one thing, estimated effects tend to be substantially smaller at the firm level; for instance, Cadot et al. (2012) find only six dollars of additional exports for one dollar of export promotion. Second, the level of detail in the decomposition of EPA activities tends to be

higher in the clinical studies than in survey-based cross-country studies, allowing for closer examination of what "treatment arms" seem to be most effective; finally, the decomposition of impacts along various margins of firm performance (extensive or intensive) is necessarily richer at the firm level. However, clinical studies have little external validity; for instance, the success of PROCHILE in fostering diversification and innovation may have to do with many features of the Chilean business and Government environment that could not be transplanted easily.

In sum, as Rodrik (2008) put it, there is an inescapable trade-off between "internal validity" (the ability to identify impact effects net of confounding influences), which improves as one goes from cross-country studies to impact evaluations, and "external validity" (the ability to draw general policy propositions from evaluation results), which may well worsen.

6. Concluding remarks

A number of observations come out of this brief survey of the issues around the potential for applying impact-evaluation techniques to trade interventions.

First, although IE is "a-theoretic", most of the practical IE literature pays at least lip service to the need for evaluation to be backed by some sort of "theory of change" (see e.g. Gertler et al. 2011). The literature we surveyed in the first part of this paper provides strong evidence of a two-step causal link from infrastructure to trade costs and from trade costs to trade performance, based on a proven empirical framework (the gravity equation) and its theoretic foundations. Thus, as far as aid to transport infrastructure is concerned, the "theory of change" is there.

In practice, aid to "hard" infrastructure often plays a twin role. Apart from its direct effect on trade costs, it also provides a hook to start or maintain dialogue with recipient governments on policy reforms, e.g. in terms of regulation of related services (trucking, maritime transport etc.) or even on broader agendas (privatization and competition issues). How much donors actually use this leverage effect of infrastructure investments varies, depending on the depth of their dialogue and their own economic sophistication. But in this "soft" area as well, the theory of change is there, as the IO-and-trade literature has long established the inter-relationship between trade performance and regulatory/competition policy (the so-called "behind-the-border" agenda).

In both cases (hard and soft infrastructure) the causal links from policy intervention to export performance are there but they are non-trivial. Thus, it is fair to say that applying IE in those areas is not what critics have derided in some IEs—to make it a bit of a caricature, counting if treating people makes them less sick.

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, as suggested by Hoekman and Wilson in their Busan paper (Hoekman Wilson 2010), 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. There is clearly a need for further thinking on this issue.

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 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

Abe, K. and J.S. Wilson (2009), "Weathering the Storm: Investing in Port Infrastructure to Lower Trade Costs in East Asia", World Bank Policy Research Working Paper 4911; Washington, DC: The World Bank.

Adam, C.S. and D.L. Bevan (2006), "Aid and the Supply Side: Public Investment, Export Performance, and Dutch Disease in Low-Income Countries", *World Bank Economic Review*, 20, 2, 261-90.

Alesina, Alberto and Dollar, David (2000). "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5(1): 33-63.

Alesina, Alberto and Weder, Beatrice (2002). "Do Corrupt Governments Receive Less Foreign Aid?" *American Economic Review* 92(4): 1126-37 (September).

Alvarez, R., and G. Crespi (2000). "Exporter Performance and promotion Instruments: Chilean Emprical Evidence"; *Estudios de Economia* 27, Universidad de Chile.

Amjadi, A. and A. Yeats (1995), 'Have Transport Costs Contributed to the Relative Decline of Sub-Saharan African Exports?', Policy Research Working Paper Series 1559, World Bank, Washington, DC.

Baier, Scott, and J. H. Bergstrand (2009), "Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation"; *Journal of International Economics* 77(1), 77-85.

Balchin, P. and S. Edwards (2008), 'Trade-Related Business Climate and Manufacturing Export Performance in Africa: A Firm Level Analysis', mimeo.

Balchin, P. and S. Edwards (2008), 'Trade-Related Business Climate and Manufacturing Export Performance in Africa: A Firm Level Analysis', mimeo.

Banerjee, Abhijit, and E. Duflo (2011*), Poor Economics*; New York, NY: PublicAffairs.

Bennathan, Ezra; Luis Escobar, and George Panagakos (1989), "Deregulation of Shipping: What is to be Learned from Chile", Policy Research Working Paper 67; Washington, DC: The World Bank.

Bernard, A. and B. Jensen (2004). "Why Some Firms Export"; *Review of Economics and Statistics* 86, 561-569.

—, Redding, S., and P. Schott (2007). "Firms in International Trade"; *Journal of Economic Perspectives* 21, 105-130.

Berger, D, W. Easterly, N. Nunn and S. Satyanath (2013) "Commercial Imperialism? Political Influence and Trade During the Cold War", *American Economic Review*, forthcoming

Bloningen, B. and W. Wilson (2008) "Port Efficiency and Trade Flows" *Review of International Economics*, 16(1), 12-36

Brenton, P. and E. von Uexkuhl (2009). "Product-Specific Technical Assistance for Exports—Has it Been Effective?" *Journal of International Trade and Economic Development* 18, 235-254.

Burnside, Craig and Dollar, David (2000). "Aid, Policies and Growth", *American Economic Review* 90(4): 847-868 (September).

Buys, Piet, Uwe Deichmann, and David Wheeler (2006), 'Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa', Policy Research Working Paper 4097; Washington, DC: The World Bank.

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

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,

Cadot, Olivier; Ana Fernandes, Julien Gourdon, and Aadityia Mattoo (2012), "Does Export Promotion Work? New Evidence from Tunisia"; mimeo, The World Bank.

— and R. Newfarmer (2012), "Does it Work? Aid for Trade through the Prism of Evaluation"; Bridges 10, http://ictsd.org/i/a4t/115387/.

Cantens, Thomas, G. Raballand, S. Bilangna and M. Djeuwo (2011), "Reforming Customs by Measuring Performance: A Cameroon Case Study"; in O. Cadot, A. Fernandes, J. Gourdon and A. Mattoo, eds., *Where to spend the next million: Applying impact evaluation to trade*; London/Washington, DC: CEPR/The World Bank, pp. 183-206.

Clemens, Michael, Steven Radelet, and Rikhil Bhavnani. 2004. "Counting Chickens When They Hatch: The Short Term Effect of Aid on Growth" CGD Working Paper 44. Washington, DC: Center for Global Development.

Coe, David T., and Alexander W. Hoffmaister. 1998. "North-South Trade: Is Africa Unusual?" Working Paper of the International Monetary Fund, Washington, D.C.

Collier, Paul and Dollar, David (2002). "Aid Allocation and Poverty Reduction," *European Economic Review* 46(8): 1475–1500.

Collier, Paul, and Jan Willem Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature,* 37:64-111.

Collier, Paul. 1995. "The Marginalization of Africa," *International Labour Review* 134(4-5):541-57.

Coulibaly, Souleymane and Lionel Fontagné, 2006, "South-South Trade: Geography Matters", *Journal of African Economies,* 15(2), 313-341.

Djankov, Simeon, Caroline Freund, and Cong S. Pham (2010), 'Trading on Time', *Review of Economics and Statistics*,92(1), 92(1), 166-73.

Freund, C. and Rocha (2011) What Constrains Africa's Exports?" *World Bank Economic Review*

Dollar, David and Levin, Victoria (2004). "The Increasing Selectivity of Foreign Aid, 1984-2002," Policy Research Working Paper No. 3299, World Bank.

Duval, Y. (2006), "Costs and Benefits of Implementing Trade Facilitations Measures under Negotiations at the WTO: an Exploratory Survey", Working Paper Series No. 3 (Asian-Pacific

Research and Training Network on Trade).

Estache, Antonio, and M. Fay (2007), "Current debates on infrastructure policy"; Policy Research Working Paper 4410; Washington, DC: The World Bank.

Estevadeordal, Antoni, and A. Taylor (2009), Is the Washington Consensus Dead? Growth, Openness, and the Great Liberation, 1970s – 2000s; CEPR Discussion Paper 6942.

Ferro, E., A. Portugal-Perez, and J. Wilson (2011). "Aid-for-Trade and Export Performance: The Case of Aid in Services," chapter 8 in O. Cadot et al. eds. pp. 207-218.

Ferro, E., A. Portugal-Perez, and J. Wilson (2012). "Aid to the Services Sector: Does it Affect Manufacturing Exports?"

Fink, C., A. Mattoo and C. I. Neagu (2002a), 'Trade in International Maritime Services: How Much Does Policy Matter?', *World Bank Economic Review*, 16, 1, 81–108.

Folletti, L., and R. Newfarmer (2011) "Aid for Trade 2011: Results emerging from the Case Stories", OECD and WTO

Foroutan, Faezeh, and Lant Pritchett. 1993. " Intra-Sub-Saharan African Trade: Is It Too Little?" *Journal of African Economies* 2(1):74-105.

Francois, Joseph and Miriam Manchin (2007), 'Institutions, Infrastructure, and Trade', Policy Research Working Paper No. 4152, World Bank, Washington, DC.

Frankel, Jeffrey A. 1997. Regional Trading Blocs in the World Economic System. Washington, D.C.: Institute for International Economics.

Freund, C.L. and D. Weinhold (2004), „The Effect of the Internet on International Trade", Journal of International Economics, 62, 1, 171-89.

Gamberoni, E. And R. Newfarmer (2009), "Aid for Trade Matching Potential Demand and Supply", Policy Research Working Paper No. 4991 (The World Bank).

Gertler, Paul; S. Martinez, P. Premand, L. Rawlings, C. Vermeersch (2011), Impact evaluation in practice; Washington, DC: The World Bank.

Girma, S., Gong, Y., Görg, H., and Z. Yu (2009). "Can Production Subsidies Explain China's Export Performance? Evidence from Firm-level Data"; *Scandinavian Journal of Economics* 111, 863-891.

Görg, H., Henry, M., and E. Strobl (2008). "Grant Support and Exporting Activity"; *Review of Economics and Statistics* 90, 168-174.

Helble, M., C. Mann and J.S. Wilson (2009), "Aid for Trade Facilitation", Policy Research Working Paper No. 5064 (The World Bank).

Helble, Matthias, Ben Shepherd, and John S. Wilson (2007), Transparency and Trade Facilitation

in the Asia Pacific: Estimating the Gains from Reform, Washington, DC: World Bank.

Hoekman, Bernard (2008). Helping Developing Countries Benefit from Trade Opportunities. In Aid for Trade and Development, ed. Dominique Njinkeu and Hugo Cameron, 27-45. Cambridge: Cambridge University Press.

Hoekman, Bernard and Alessandro Nicita (2008), 'Trade Policy, Trade Costs, and Developing Country Trade', Policy Research Working Paper No. 4797, World Bank, Washington, DC.

— and J. Wilson (2010), "Aid for Trade: Building on Progress to Date"; paper prepared for the Korea-World Bank High Level Conference on Post-Crisis Growth and Development, Busan, June 3-4, 2010.

Hummels, D., V. Lugovskyy and A. Skiba (2009), "The Trade Reducing Effects of Market Power in International Shipping", *Journal of Development Economics*, 89, 1, 84-97.

Hummels, David (1998), "Toward a geography of transport costs"; mimeo, University of Chicago.

Hummels, D. and V. Lugovskky, (2006), "Are Matched Partner Trade Statistics a Usable Measure of Transportation Costs?", *Review of International Economics*, 14, 69-86

Hummels, David (1999) , "Have International Transport Costs Declined?"; mimeo, University of Chicago.

Keesing, Donald B. and Andrew Singer (1991), "Development assistance gone wrong: failures in services to promote and support manufactured exports", in Paul Hogan, Donald Keesing, and Andrew Singer eds., The Role of Support Services In Expanding Manufactured Exports in Developing Countries, Economic Development Institute; Washington, DC: The World Bank.

Khandker, Shahidur; G. Koolwal and S. Samad (2009), Handbook on impact evaluation: Quantitative methods and practices; Washington, DC: The World Bank.

Lederman, D., Olarreaga, M., and L. Payton (2010). "Export Promotion Agencies Revisited"; *Journal of Development Economics* 91, 257-265.

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

Moneta, C. 1959. "The Estimation of Transport Costs in International Trade." *Journal of Political Economy* 67:41-58.

Nelson, D. and S. Silva (2008). "Does Aid Cause Trade? Evidence from an Asymmetric Gravity Model." University of Nottingham research paper 2008/21.

Neumayer, Eric (2003a). "Is Respect for Human Rights Rewarded? An Analysis of Total Bilateral and Multilateral Aid Flows," *Human Rights Quarterly* 25(2): 510-27.

Neumayer, Eric (2003b). "Do Human Rights Matter in Bilateral Aid Allocation? A Quantitative Analysis of 21 Donor Countries," *Social Science Quarterly*, 84(3): 650-66. 34

Njinkeu Dominique, John S. Wilson, and Bruno Powo Fosso (2008), ' Intra-Africa Trade Constraints : The Impact of Trade Facilitation', mimeo, World Bank.

Njinkeu Dominique, John S. Wilson, and Bruno Powo Fosso (2008), ' Intra-Africa Trade Constraints : The Impact of Trade Facilitation', mimeo, World Bank.

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

Inquiry

OECD (2009a). How to Evaluate Aid for Trade? A Scoping Note, Paris.

OECD (2011) " Strenghtening Accountability in Aid for Trade" OECD

Osei, R., O. Morrissey, and T. Lloyd (2004). "The Nature of Aid and Trade Relationships." *European Journal of Development Research* 16, 354-374.

Page, Sheila (2007). The Potential Impact of the Aid for Trade Initiative, UNCTAD G-24 Discussion Paper Series No. 45.

Page, Sheila (2008). Aid for Trade: A New Issue in the WTO. In Aid for Trade and Development, ed. Dominique Njinkeu and Hugo Cameron, 131-150. Cambridge: Cambridge University Press.

Portugal-Perez, A. and J.S. Wilson (2008), "Trade Costs in Africa: Barriers and Opportunities for Reform", Policy Research Working Paper No. 4619 (The World Bank).

Raballand, G., C. Kunaka, and B. Giersing (2008), 'The Impact of Regional liberalization and Harmonization in Road Transport Services: A Focus on Zambia and Lessons for Landlocked Countries', Policy Research Working Paper 3840;

Raballand, Gaël; Rebecca Thornton, Dean Yang, Jessica Goldberg, Niall Keleher, Annika Müller (2011), "Are Rural Road Investments Alone Sufficient to Generate Transport Flows? Lessons from a Randomized Experiment in Rural Malawi and Policy Implications", World Bank Africa Region Transport Department working paper 5535, Washington, DC: The World Bank.

Rajan, R. and A. Subramanian (2008). "Aid and Growth: What Does the Cross-Country Evidence Really Show?" *Review of Economics and Statistics* 90, 643-665.

Redding, S. and A.J. Venables (2003), "South-East Asian Export Performance: External Market Access and Internal Supply Capacity", *Journal of the Japanese and International Economies*, 17, 4, 404-431.

Redding, S. and A.J. Venables (2004), "Geography and Export performance: External Market Access and Internal Supply Capacity", in R.E. Baldwin and L.A Winters (eds), Challenges to Globalization: Analyzing the Economics (Chicago: University of Chicago Press).

Rodrik, Dani (2008), "The New Development Economics: We Shall Experiment, But How Shall We Learn?"; Mimeo, Kennedy School of Government, Harvard University.

Shepherd, Ben(2006), 'Transport Costs and Trade over Time', in OECD Report of the One Hundred and Thirtieth Round Table on Transport, Round Table 130, Transport and International Trade.

Shepherd, Ben, and J. Wilson (2006), "Road infrastructure in Europe and central asia: does network quality affect trade?" Policy Research Working Paper 4104; Washington, DC: The World Bank.

Suwa-Eisenmann, Akiko and Verdier, Thierry (2007). "Aid and Trade*," Oxford Review of Economic Policy* 23(3): 481-507.

Tadasse, Badassa and Fayissa, Bichaka (2009). "Determinants of the Allocation of U.S. Aid for Trade," Working Papers 200901, Middle Tennessee State University, Department of Economics

and Finance.

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 (2012), "Aid for trade effectiveness: complementarities with economic integration"; mimeo, FERDI.

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.

Wilmeister, G. and J. Hoffman (2008) "Liner Shipping Connectivity and Port Infrastructure as Determinants of Freight Rates in the Caribbean", *Maritime Economics and Logistics* 10, 131-151

Wilson, John S., Catherine Mann, and Tsunehiro Otsuki (2003), "Trade Facilitation and Economic Development : A New Approach to Measuring the Impact"; *World Bank Economic Review*, 17(3) 67–389.

World Bank (2007), 'Logistic Performance Index database', available at http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTTRANSPORT/EXTTLF/0,,contentMDK:21514122~menuPK: 3875957~pagePK: 210058~piPK:210062~theSitePK:515434,00.html. 1025

World Bank (2007), 'Logistic Performance Index database', available at http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTTRANSPORT/EXTTLF/0,,contentMDK:21514122~menuPK: 3875957~pagePK: 210058~piPK:210062~theSitePK:515434,00.html. 1025

——(2008), 'Doing Business database', available at http://www.doingbusiness.org/.

——(2008), 'Doing Business database', available at http://www.doingbusiness.org/.

Annex

Aid Classification in the OECD data base

The OECD Creditor Reporting System (CRS) database covers around 90% of all ODA. It tracks aid commitments and disbursements, and provides comparable data over time and across countries. It is recognized as the most complete and wide-ranging available data source for tracking global aid-for-trade flows in that it was already firmly established and in use, saving time and resources. However, it entailed some loss of detailed information about trade-related technical assistance and trade development activities which were collected by the more specialized joint OECD-WTO Trade Capacity Building Database (TCBD). Reporting in that data base was discontinued in 2007. Several modifications have been made to the CRS categories to adapt it to AFT (e.g. the new CRS category "trade related adjustment" and the "trade development marker" introduced in 2008 covering 2007 flows but not yet used). Table A-1 shows how the AFT categories are matched to the CRS categories.

Table A.1
AFT categories in the CRS Data base

Aid for Trade categories	Proxies in CRS data
Technical assistance for trade policy and regulations (e.g. helping countries to develop trade strategies, negotiate trade agreements, and implement their outcomes)	Five sub-categories: (a) trade policy and administrative management; (b) trade facilitation; (c) regional trade agreements; (d) multilateral trade negotiations; and, (e) trade education/training.
Trade-related infrastructure (e.g. building roads, ports, and telecommunications networks to connect domestic markets to the global economy)	Proxied in the CRS by data under the heading "economic infrastructure". This heading covers data on aid for communications, energy, transport and storage. To know how close the CRS proxies are the CRS data must be compared with donors' knowledge of the specific features of their aid to infrastructure .
Productive capacity building, including trade development (e.g. supporting the private sector to exploit their comparative advantages and diversify their exports)	Data on commitments of aid for productive capacity building exist under the CRS category "building productive capacity" and covers Agriculture, Forestry, Fishing, Mining, Industry and Services sectors
Trade-related adjustment (e.g. helping developing countries with the costs associated with trade liberalisation, such as tariff reductions, preference erosion, or declining terms of trade)	This category in CRS identifies contributions to developing country budgets to assist in the implementation of trade reforms and adjustments to trade policy measures by other countries
Other trade-related needs, if identified as trade-related development priorities in partner countries' national development strategies	The CRS covers all ODA, but only those activities reported under the above four categories will be identified as aid for trade. Data on "other trade-related needs" cannot be gleaned from the CRS.

In this paper, we use this new AFT database from CRS. However, with this database the only way to obtain data for commitments prior to 1995 and for disbursements prior to 2002 is to download the initial CRS dataset that does not account for the recent modifications mentioned above. To be closer to our trade cost approach in Figure 1 we modify this AFT structure. These modifications are shown in figure A2. Finally, table A3 shows the share of total aid devoted to those components in 2002 and 2010 as well as the name of specific programs.

Table A.2
Matching AFT categories to the classification in Figure 1

Aid for Trade categories	Classification in figure 1 and corresponding CRS data
Technical assistance for trade policy and regulations	Trade Policy in CRS: (a) trade policy and administrative management; (c) regional trade agreements; (d) multilateral trade negotiations; and, (e) trade education/training. And trade-related adjustments
	Border related Cost in CRS (b) trade facilitation;
Trade-related infrastructure (e.g. building roads, ports, and telecommunications networks to connect domestic markets to the global economy)	Hard Infrastructure In CRS: aid for transport and storage. aid for energy is dropped Soft Infrastructure In CRS: aid for communications
Productive capacity building	Regulatory Policies In CRS aid for Banking Financial Services and Business Services and Aid for Production sectors are dropped
Trade-related adjustment (e.g. helping developing countries with the costs associated with trade liberalisation, such as tariff reductions, preference erosion, or declining terms of trade)	Moved to Trade Policy
Other trade-related needs, if identified as trade-related development priorities in partner countries' national development strategies	None.

Source :Authors' compiled from the CRS website
AFT categories from XXX

Table A.3
Aid for Trade by categories: 2002 and 2010

	2002	2010
1. Hard Infrastructure	43.2	52.7
1.1 Transport Infrastructure	28.7	30.2
Transport policy and administrative management	1.8	3.5
Road transport	17.5	16.8
Rail transport	6.6	7.4
Water transport	1.4	1.6
Air transport	1.3	0.9
Other	-	0.1
1.2 Energy Infrastructure	14.6	22.5
Electrical transmission/ distribution	3.8	6.0
Power generation/renewable sources	0.8	4.2
Energy policy and administrative management	2.6	2.9
Power generation/non-renewable sources	2.3	2.6
Hydro-electric power plants	1.1	1.6
Gas-fired power plants	-	1.4
Solar energy/Wind Power	0.3	1.9
Nuclear power plants	1.1	1.0
Other	2.6	1.0
2. Behind the borders Policies	15.7	13.9
2.1 Soft Infrastructure	1.9	1.6
Communications policy and administrative management	0.4	0.3
Telecommunications	1.1	0.4
Radio/television/print media	0.4	0.4
Information and communication technology (ICT)	-	0.6
2.2 Banking Financial Services	9.4	7.5
Financial policy and administrative management	3.5	1.9
Monetary institutions	-	0.1
Formal sector financial intermediaries	4.0	3.4
Informal/semi-formal financial intermediaries	1.8	2.1
Education/training in banking and financial services	-	0.1
2.3 Business Sertvices	4.3	4.7
Business support services and institutions	4.1	4.5
Privatisation	0.2	0.2
3. Trade Policies	4.8	2.8
Trade policy and administrative management	4.3	1.7
Regional trade agreements (RTAs)	0.2	0.3
Multilateral trade negotiations	-	0.1
Trade-related adjustment	-	0.2
Trade education/training	0.1	0.1
Tourism policy and administrative management	0.2	0.5

Table A3 (continued)

	2002	2010
4. Trade facilitation	-	1.0
5. Productive capacity building	36.3	29.7
5.1 Primary	24.6	24.9
Agricultural development	3.8	5.0
Agricultural policy and administrative management	2.7	5.0
Agricultural water resources	2.7	2.4
Agricultural research	1.9	1.6
Agricultural alternative development	0.1	1.0
Other Agricultural	9.3	5.4
Forestry	2.6	3.6
Fishing	1.5	0.9
5.2 Industry	7.3	4.3
Small and medium-sized enterprises (SME) development	0.8	2.4
Industrial development	4.5	0.5
Agro-industries	0.3	0.4
Industrial policy and administrative management	1.0	0.2
Engineering	-	0.2
Technological research and development	0.1	0.2
Other	0.6	0.3
5.3 Mining	4.4	0.5
Mineral/mining policy and administrative management	2.2	0.3
Mineral prospection and exploration	-	0.1
Oil and gas	1.9	0.1
Other	0.2	-

Source: Authors' compiled from the CRS website according to classification in table A2.