HAL open science

The Distance Effect and the Regionalization of the Trade of Low-Income Countries

Céline Carrere, Jaime Melo De, John Wilson

▶ To cite this version:

Céline Carrere, Jaime Melo De, John Wilson. The Distance Effect and the Regionalization of the Trade of Low-Income Countries. 2011. halshs-00554326

HAL Id: halshs-00554326

https://shs.hal.science/halshs-00554326

Preprint submitted on 10 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CENTRE D'ÉTUDES
ET DE RECHERCHES
SUR LE DÉVELOPPEMENT
INTERNATIONAL

COMMENTS WELCOME

Document de travail de la série

Etudes et Documents

E 2009. 08

THE DISTANCE EFFECT AND THE REGIONALIZATION OF THE TRADE OF LOW-INCOME COUNTRIES

Céline Carrère[+]
CERDI – CNRS

Jaime de Melo[♣]
University of Geneva and CERDI

John Wilson[♦]
World Bank

February 2009

[+] CERDI (University of Auvergne, France) and CNRS
[♣]University of Geneva, CERDI and CEPR
[♦]The World Bank

*Thanks to participants at seminars at the World Bank and CERDI. This paper has benefited from support from DFID.

Abstract

The "distance effect" measuring the elasticity of trade flows to distance has been to be rising since the early 1970s in a host of studies based on the gravity model, leading observers to call it the "distance puzzle". We review the evidence and explanations. Using an extensive data set of 124 countries over the period 1970-2005, we confirm the existence of this puzzle and identify that it only applies to poor countries (the bottom third in per capita income terms in our sample—i.e. the low-income countries according to the World Bank classification, 2006). We show that this group has intensified trade with closer partners and have chosen new partners that are closer than existing partners, leading to a regionalization of their trade at both extensive and intensive margins (regionalization of trade is absent for the other countries). Combining several methods on cross-section and panel estimates of the gravity equation, we estimate that low-income countries exhibit a significant rising distance effect on their trade around 15% between 1970 and 2006 while there is no more distance "puzzle" for trade within richer countries (the top third in per capita income terms in our sample). We dispose of previous explanations of the puzzle, and note that this regionalization could well be a reflection of both increased integration of this group of countries in the world economy or a greater marginalization.

JEL Classification: F10; F40

Keywords: International Trade; Gravity Model; Distance Effect

1. INTRODUCTION

There is a widespread perception that the current wave of globalization, much like the first, should have led to the "death of distance". Its importance for developing countries cannot be underestimated since under a broad range of models, the magnitude of distance effects determines the wage gap with richer countries and the ability to attract footloose industries. In a more popular vein, as argued by Thomas Friedman in *The World is Flat* (2005), the fall in communication costs which are an integral part of overall transactions costs that are captured by distance, should provide a tremendous opportunity for the poorer countries to integrate the world economy especially because of their backwardness and the rapid spread of reduction in these costs around the world. With quasi-costless communication, outsourcing will increase and producers in remote developing countries will now be able to supply far-away Northern markets for fashion and other differentiated products with relatively short shelf-lives.

Under this popular interpretation of the "death of distance" scenario, ceteris paribus, the average distance of trade for poorer countries should increase (as lower transport costs would open more distant markets). Yet, no visible increasing trend in the average distance of trade has been detected in the data over the last thirty years for the poorest. Surprising at it may seem, this is coherent with the high trade costs reported by Anderson and Van Wincoop (AvW, 2004) in their survey. It is also consistent with the "distance puzzle" which suggests that the burden of distance on trade may have, in fact, been increasing. In terms of the gravity literature used to estimate trade costs, a reduction in trade costs should imply a smaller "distance effect", i.e. a declining value (in absolute terms) of the elasticity of trade to distance, θ. The challenge then is how to reconcile technology driven reduction in trade costs with a non-shrinking effect of distance in the large literature estimating trade costs from data on bilateral trade.

This puzzle is the subject of this paper. Although we provide several estimates of θ, someone of which are more plausible than previous ones, this is not our main concern (Grossman, 1998 and AvW, 2004 object to its high estimated value).[1] Instead, we focus on the causes for the persistent finding that it is increasing with time. We focus on the poorest countries because it is precisely this group of countries that should benefit most

[1] See the discussion of Grossman's objection in Anderson and Van Wincoop (2004, p. 729). Taking into account the quality of infrastructure and the choice of mode of transport (e.g. air, sea) for a cross-section sample of Asian countries, Shepherd and Wilson (2006) obtain a lower estimate $\tilde{\theta} = -0.35$.

from a flatter world. This focus is also motivated by our earlier work (Brun et al., 2005) where we found that an increasing 'burden' of distance was restricted to poor countries and conjectured that they may have been marginalized by the current wave of globalization.

Section 2 reviews the explanations for this finding: composition effects, sample selection, econometric methods and omitted variable bias. This preliminary exploration, in part based on a meta-analysis on 103 published paper (based on the Disdier and Head (2008)'s database), leads us to suggest that the puzzle mainly holds for developing countries. The rest of the paper seeks to check if this is still the case when determinants of distance-sensitive trade costs are brought into the determination of bilateral trade. Section 3 analyzes the evolution of the average distance of trade, confirming and sharpening the results from the meta-analysis. Over time, the third poorest countries in the sample of 124 countries (i.e. the low-income countries according the World Bank classification of 2006) shift, among existing trade partners, towards physically closer partners. Also, their new trading partners are closer than existing partners. No such pattern is apparent in the data for the remaining countries. These findings confirm a changing role of distance in bilateral trade. In section 4, we revisit the gravity-predicted θ elasticities. To control for as many factors as possible, and to maximize robustness, we rely on cross-section and panel formulations and use several methods to deal with zero trade flows. In all cases, a distance puzzle is revealed for the bottom third (39 countries) in the sample, leading us to conclude that trade has become regionalized for low income countries. Section 5 concludes.

2. The Distance Puzzle in the Gravity Literature

2.1. The Rising Distance Effect

While there are several approaches to estimate the impact of transport costs on the volume of trade, the great majority of estimates rely on the popular gravity model which states that the volume of bilateral trade between two countries (*i* and *j*) should be proportional to their economic size, proxied by GDP ($Y_{i(j)}$) and inversely proportional to transport costs, proxied by the distance between partners (D_{ij}). The numerous studies in the literature deliver an estimate of the elasticity of bilateral trade to distance, $\tilde{\theta}$, which is then used to predict bilateral trade volumes as a function of distance. For example, using the range of estimates in the literature, with $\tilde{\theta} = -1.4$ $[-0.7]$ doubling the distance reduces trade by 63%

[42%].[2] This range is typical of cross-section (sometimes averaged over 5-year periods) estimates of aggregate trade volumes where trade costs are given by:

$$t_{ij} = \left(D_{ij}\right)^{\rho} \prod_{m=1}^{M} \left(z_{ij}^{m}\right)^{\gamma_m} \qquad (1)$$

Where, the set z_{ij}^{m} (m=1,...,M) includes binary dummy variables (usually invariant through time, such as sharing a common border, a common language, etc.) capturing other barriers to trade than distance. These costs enter log-linearly in the "traditional" gravity equation:

$$\ln\left(M_{ij}\right) = \alpha_0 + \alpha_1 \ln\left(Y_i\right) + \alpha_2 \ln\left(Y_j\right) + \theta \ln\left(D_{ij}\right) + \sum_{m=1}^{M} \lambda_m \ln\left(z_{ij}^{m}\right) + \varepsilon_{ij} \qquad (2)$$

and the distance effect is given by the estimate $\theta = \alpha\rho$, with $\alpha < 0$ being the trade elasticity to trade costs t_{ij}. As discussed by AvW (2004), and as illustrated in the above typical estimates of $\tilde{\theta}$, these appear to be implausibly high. [3] This lead AvW to conclude that distance is in fact capturing other barriers to trade (e.g NTBs, information barriers, and contracting Costs and insecurity) not appropriately controlled for in the set of dummy variables z_{ij}^{m}. [4]

But the real "puzzle" is that estimates of $\tilde{\theta}$ coming from more recent data yield larger estimates. These results imply that distance has exerted a more powerful (negative) effect on the volume of trade in recent times. This is clear from figure 1 below reproduced from a recent meta-analysis of 1,467 elasticity estimates $\tilde{\theta}$ compiled from gravity model estimates reported in over 100 published papers (see Disdier and Head, 2008). This figure plots the elasticity estimates against time and fit a kernel smoother through the

---

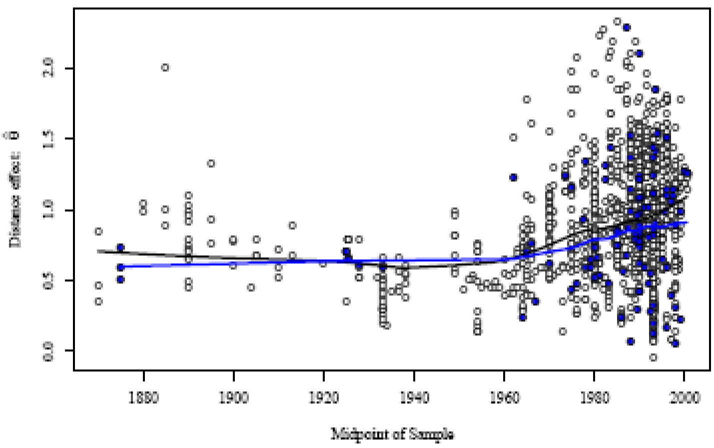
[2] The general formula is: $M_1 / M_0 = \left(D_1 / D_0\right)^{\tilde{\theta}}$

[3] In the theory-based gravity equation, the elasticity of trade to trade costs depends on the elasticity of substitution in consumption, σ according to $\alpha = (\sigma - 1)$. Since σ has to be estimated separately (as reported by Anderson and Van Wincoop, 5<σ<10), the elasticity of trade to distance will, in fact, depend on the ease with which goods can be substituted across suppliers. As we discuss below, the composition of trade would then appear to matter.

[4] Based on σ=8, they estimate that the overall border barriers to trade amount to around 50% .

data (dark line).[5] From their survey of estimates reported in figure 1 and from further analysis of the evolution of the estimates through time (see below), Disdier and Head (2008) conclude that the evolution of the distance impact on trade was fairly flat until the 1950s, but has shown a significant increase in the post-1970 data.

Figure 1. The rising Distance Effect in Gravity Models



Source: Disdier and Head (2008, figure 3, p.19).

To give an idea of the orders of magnitude suggested by the meta-analysis summarized in figure 1, distance impedes trade by 37% more since 1990 that it did from 1870 to 1969. This increasing elasticity of trade to distance had already been noticed by Frankel (1997). Earlier, Leamer and Levinsohn (1995, pp. 1387–88), reviewing the literature on international trade and distance, noted that "the effect of distance on trade patterns is not diminishing over time. Contrary to popular impression, the world is not getting dramatically smaller." This paradoxical result, now well established, is referred to as the "distance puzzle" or the "missing globalization puzzle" (Coe et al. 2007).

2.2. The Gravity model Set-up

Even though most estimates in figure 1 come from the "traditional" gravity equation in (2), it is now recognized that gravity-based estimates of changes in trade costs give more intuitive and plausible results when obtained from theory-based gravity models that point out explicitly the channels through which bilateral trade depends on relative trade costs, and indirectly, to distance. To take an example, given trade costs (partly

5 The highest R^2 estimate of each paper is shown with a solid circle, and the lighter blue lines report the associated lowess smoother estimates.

proxied by distance) will matter less for bilateral trade between New-Zealand and Australia than for bilateral trade between Greece and Switzerland because Australia and New-Zealand are further away from their other trade partners than Greece and Switzerland. A large family of trade models satisfies the conditions necessary to yield a gravity equation at the product level. [6] Here is one. Take a one sector economy with a representative consumer with CES preferences with common elasticity σ among all goods. Impose symmetry of trade costs ($t_{ij}= t_{ji}$) and assume that trade costs are proportional to trade (no economies of scale in transport). Then the delivered price includes an ad-valorem equivalent of trade costs (tariffs, NTBs, etc.). With constant returns to scale in transport and marginal cost pricing in transport $p_{ij} = p_i t_{ij} = p_i\left(1+\tau_{ij}\right)$ and trade costs enter multiplicatively as in (2). Under these assumptions, outward and inward trade costs $\tilde{P}_i$, $\tilde{P}_j$ are symmetric and the theory-based gravity equation is composed of the system of equations:

$$M_{ij} = \left(\frac{Y_i Y_j}{Y_w}\right)\left(\frac{t_{ij}}{\tilde{P}_i \tilde{P}_j}\right)^{1-\sigma} \tag{3}$$

$$\tilde{P}_i^{1-\sigma} = \sum_j \frac{Y_i}{Y_w}\left(\frac{t_{ij}}{\tilde{P}_j}\right)^{1-\sigma} \tag{4}$$

where M_{ij} is the imports of country *i* from country *j*, $Y_{i(j)}$ is the GDP of country *i (j)*, Y_w is world GDP, t_{ij} is bilateral trade costs between *i* and *j*, σ >1 is the elasticity of substitution in the CES utility function. According to (3) and (4), bilateral trade flows depend on the relative size of partners and conditionally on relative trade costs where $\tilde{P}_i$, $\tilde{P}_j$ respectively represent the inward and outward multilateral trade resistance.

A more satisfactory formulation of trade would recognize that transaction costs include several components, and that per-unit transport prices may not be equal to transport costs because of market power by transport carriers. Using disaggregated US ocean freight rates over the 1991-2004

[6] See AvW (2003, 2004). These conditions are: (i) trade separability (i.e. separability in preferences and technology as in CES technology and utility); (ii) aggregator of varieties are identical and CES across countries; (iii) trade costs are proportional to trade and may include local distribution costs, but these costs do not affect trade flows; (iv) consumer have CES preferences with a common elasticity of substitution σ across commodities. Trade costs do not depend on the quantity of trade, a strong assumption since trade costs are likely to depend on the volume of trade.

period and a cross-section of Latin American freight rates, Hummels et al. (2009) find that ocean-carrier markups are particularly sensitive to tariffs in Latin America and that, jointly with product characteristics, they explain an order of magnitude more of the variation in shipping prices than distance.[7] Thus changes in trade policy and in the degree of competition in shipping will change the ad-valorem equivalent of trade costs lumped here for convenience under the term τ_{ij}. The reduced-form distance-dependent trade cost function would read:

$$t_{ij} = \left(1+\tau_{ij}\right)\left(D_{ij}\right)^{\rho} \prod_{m=1}^{M} \left(z_{ij}^{m}\right)^{\gamma_m} \tag{5}$$

where the ad-valorem equivalent of trade costs includes all border trade costs, depend on product characteristics and on the market characteristics of the transport sector.[8] In practice, the functional form of trade costs t_{ij} is in fact given by (1). Substituting (1) into (3) and (4), the estimated equation in a cross-section setting becomes:

$$\begin{aligned} \ln\left(M_{ij}\right) = & -\ln\left(Y_w\right)+\ln\left(Y_i\right)+\ln\left(Y_j\right)-(\sigma-1)\rho\ln\left(D_{ij}\right)+\sum_{m=1}^{M}(\sigma-1)\gamma_m\ln\left(z_{ij}^{m}\right) \\ & +(\sigma-1)\ln\left(\tilde{P}_i\right)+(\sigma-1)\ln\left(\tilde{P}_j\right)+\varepsilon_{ij} \end{aligned} \tag{6}$$

In this formulation the income coefficient terms are unity, remoteness terms sometimes included in the estimation are derived directly from theory (and called multilateral resistance terms) and the “distance effect” becomes $\theta = (\sigma-1)\rho$, the distance elasticity depending on composition effects. Thus a country that would trade mostly homogenous goods with close substitutes would face very small trade costs and the gravity model would not be useful to learn about trade costs in those circumstances.[9] Estimates of (6) can be computed from several methods, but most of them are obtained by the inclusion of country fixed-effects which is addition to

[7] They find that few carriers and high tariffs contribute to the significantly higher shipping prices facing developing countries and estimate that a 1% reduction in the tariff reduces the shipping price by 1.2% to 2.1% .

[8] Even though we follow the literature and use the multiplicative form of the trade cost function, it has been criticized as it implies that the marginal effect of a change in one cost depends on all other costs. Hummels (1999) suggests the alternative additive trade cost function $t_{ij} = \left[\left(f_{ij}+\tau_{ij}\right)\left(D_{ij}\right)\right]^{\rho} + \sum_{m=1}^{M}\gamma_m z_{ij}^{m}$ where f_{ij} is the freight rate.

[9] The poor performance of the gravity model for trade among low-income countries is well-known. Feenstra (2004, chp.5) attributes this poor performance to the non-fulfillment of the critical condition of specialization in different commodities for low-income countries.

producing unbiased estimates, avoids the measurement errors inherent in the use of price indexes. About one quarter of the estimates reported by Disdier and Head (2008) includes these country fixed-effects. As before, a "distance puzzle" obtains when the distance effect, $\hat{\theta}$, takes larger values when estimated from more recent data as exemplified in figure 1.

Economic and econometric arguments have been advanced to explain the presence of the puzzle. Unfortunately, Disdier and Head (2008) cannot test for each effect separately from the estimates reported in their meta-analysis. For convenience, we categorize these arguments under the headings of sample composition, zeroes in the data and incorporation of multilateral resistance and look for the sensitivity in the estimates of $\hat{\theta}$ to these three set of controls, using the meta-analysis dataset of Disdier and Head (2008).

2.3. Sensitivity of Distance Elasticity estimates in existing empirical literature

Composition Effects. Composition effects appearing through the elasticity of substitution at the product level, have been invoked most. [10] Two recent studies shed some light. Estimating the distance elasticity of bilateral trade for 700 manufactured products in a sample including developing and industrialized countries, Berthelon and Freund (2008) find no evidence that changes in the composition of trade across manufactured commodities accounts for the distance puzzle (but they give some evidence that for 40% of industries distance became more important). However, compositional changes could take place between broader categories of products. In this vein, Melitz (2007) finds supporting evidence for the argument that there might have been a shift in trade patterns from comparative-advantage-based to intra-industry trade in differentiated products with intra-industry trade mostly among North-North countries that share similar characteristics. Distance has a positive impact on comparative-advantage-based (Ricardian) trade since differences in endowments /productivities are positively correlated with distance while it has a negative impact for trade in differentiated products. Then, if the share of trade based on comparative advantage decreases (which has certainly been the case if one considers the evolution of the share of agricultural trade in total trade), the negative impact of distance on trade will increase mechanically. In a sample including developed and developing countries, Melitz (2007) shows that when he introduces the

[10] Disdier and Head (2008) discuss three channels that will yield different estimates of θ in theory-based gravity models: differences in σ across products, differences in the response of trade costs to distance across goods ρ and differences in productivity across firms (see also the discussion in Anderson and Van Wincoop, 2004, p.726).

difference in latitude (as a proxy for Ricardian trade), the elasticity estimate of trade to distance falls by half when estimated in several cross-sections over the period 1970 to 1995.[11]

Composition effects may also be at work through omitted variable bias. Consider for example the impact that the quality and quantity of social and physical infrastructure may have on trade costs that may be captured in the elasticity of trade to distance. Trade costs may be higher in countries with poor-quality institutions (institutions have been found to be persistent and to change little through relatively long time periods). Then falling communication costs would result in a smaller decrease of trade costs in countries with low-quality social infrastructure. Francois and Manchin (2006) find supporting cross-sectional evidence. Likewise, when they introduce a proxy for contractual enforcement and corruption in the trade cost function, Anderson and Marcouiller (2002) find that the implied tax equivalent of relatively low-quality institutions is 16%. Along similar lines, Aidt and Gassebner (2008) find that autocratic states trade less. While neither finding deals directly with the elasticity of trade to distance, nor with its evolution, they suggest that omitted variable bias could have a systematic impact on the evolution of the elasticity of trade to distance through their impact on trade costs.

Physical infrastructure could also play an important role as first shown by Limao and Venables (2001). Brun et al. (2005) estimated an "augmented" gravity equation incorporating a time-varying indicator of the quality of physical infrastructure. [12] The quality of physical infrastructure has also been brought to light in recent estimates that incorporate indicators of the quality of road infrastructure (Buys et al., 2006 for Africa and Shepherd and Wilson, 2006 for Central Asia).

The characteristics of trade costs could also contribute towards explaining the puzzle. Since international trade involves fixed costs (see the discussion of evidence in AvW 2004), if technological progress in shipping has been relatively slow in comparison to technical progress in the rest of the economy, then the puzzle could show up in the data through an increase in transport costs as a fraction of average production costs. This is the interpretation of Estevadeordal et al. (2003) who estimated the

[11] For reference, when we control for the composition of export by the share of primary products in total trade, we do not find any effect on the estimated value of θ.

[12] With this formulation in random-effects estimation over the period 1962-96, Brun et al. (2005) find falling values through time for $\tilde{\theta}$ for trade between countries in the richest tercile in the sample, but they find that the distance puzzle persists for trade between the poorest-tercile countries in the richest tercile. Shepherd and Wilson (2006) also obtain evidence that the quality of infrastructure matters for the volume of bilateral trade.

elasticity of trade to distance for 1913, 1928 and 1938. Brun et al. (2005) and Carrère and Schiff (2005) have also suggested this interpretation: the elasticity of transport costs with respect to distance could increase if the fixed cost component (dwell costs such as port storage costs, loading and unloading costs, time in transit, tariffs on imports, etc.) were falling sufficiently faster than the variable component (e.g. fuel costs, costs of manning and leasing ships). Brun et al. (2005) find that the puzzle holds for developing, but not for developed countries. Finally, Hummels (2001) and Deardorff (2003) suggest that the influence of time on trade is increasing because of greater use of just-in time production. Then this would show up as rising distance costs.

Handling zeroes in the data. Recent contributions have explored the treatment of zeroes in the data and the handling of the multilateral resistance terms. As argued by Santos Silva and Tenreyro (2006), Martin and Pham (2007) or Eaton and Tamura (1994), ignoring the zero-trade data can severely bias gravity equation estimates. Felbermayr and Kohler (2006) show that standard OLS estimates on the sample of positive traders will yield downwards- biased estimates of the distance coefficient on early data (as zero trade flows due to high trade costs are not taken into account) while more recent estimates (with less zero trade flows) are closer to the “true” values. In other words, if zero trade flows are positively correlated with large distance (which is clearly the case as discussed below), then ignoring zero trade flows when estimating the gravity equation can generate an “artificial” or spurious distance puzzle.

Relatedly, omitting the multilateral terms when estimating (6) generates a bias in the estimation of $\theta = (\sigma - 1)\rho$ since the bilateral distance is correlated with these multilateral terms that are left in the error term ε_{ij} (see the discussion in AvW, 2004, page 714).

Finally is the issue of the appropriate functional form for the trade cost function. Coe et al. (2007) find declining distance effects when they specify the gravity equation with an additive error term and estimate it using nonlinear least squares. However, as emphasized by Anderson and Van Wincoop (2004), this is not clear why such estimation would resolve the puzzle. Moreover, using Monte Carlo simulations, Santos Silva and Tenreyro (2006) find that the nonlinear least squares estimator performs very poorly.

Disdier and Head (2008) started exploring these competing explanations in empirical literature by estimating the following correlates of the distance coefficient in their sample of estimates:

$$\tilde{\theta}_{ij} = \alpha_0 + \alpha_1 D^{70-79} + \alpha_2 D^{80-89} + \alpha_3 D^{90-99} + \sum_m \beta_m x_{ij}^m + u_i + e_{ij} \qquad (7)$$

where $\tilde{\theta}_{ij}$ is the *j*th distance coefficient reported in study *i*, the D variables are dummies taking values of 1 when the midyear of the sample used to estimate the *j*th distance coefficient in study *i* is in the 70s, 80s and 90s respectively and x_{ij}^m is a set of dummy controls. These dummies control for the presence of developing countries in the sample, for a correction for the zero trade flows, for the use of a country fixed effect, for disaggregated data, etc. (see their table 2 in 44). Finally, the u_i are random effects. In (7), positive estimates for α_1, α_2 and α_3 represents the additional distance effect in, respectively, [1970-1979], [1980-1989] and [1990-1999] compared to pre-1970, once controlled for systematic differences in the attributes of the studies through the *x* vector.

Their results show that using a sample restricted to developing countries increases significantly the distance elasticity by 0.44 percentage point in the random-effects specification (see their table 2 col. 4 page 44). Likewise, they find that incorporating zero trade flows in the sample or introducing country fixed effects (i.e. specifying a gravity equation consistent with theory) increases the distance coefficient by 0.08 and 0.14 percentage points respectively. However, even after controlling for all these aspects of the estimates that could "artificially" create the distance puzzle observed in figure 1, the increasing distance effect after 1970 remains. In sum, the meta-analysis persists in showing a rising estimate of $\tilde{\theta}$ across samples, specifications and econometric methods.

To see if these influences vary with time, we extend their exploration by interacting the controls, x_{ij}^m with time dummies, i.e. we estimate: [13]

$$\begin{aligned}\tilde{\theta}_{ij} &= \alpha_0 + \alpha_1 D^{70-79} + \alpha_2 D^{80-89} + \alpha_3 D^{90-99} \\ &\quad + \alpha_1^m \left(D^{70-79} x_{ij}^m\right) + \alpha_2^m \left(D^{80-89} x_{ij}^m\right) + \alpha_3^m \left(D^{90-99} x_{ij}^m\right) + \sum_m \beta_m x_{ij}^m + u_i + e_{ij}\end{aligned} \qquad (8)$$

We report here results for the three dummy variables that identify the presence of: (i) developing countries; (ii) corrections for the zero trade flows; (iii) controls for the multilateral trade resistance terms suggested by theory i.e. by including either remoteness variables or country fixed effects.

[13] We thank Anne-Célia Disdier and Keith Head for sharing their database.

Coefficients of interest are reported in figures 2a-2c. In general, the coefficient estimates do not vary across time except for studies focusing on developing countries where a very strong distance puzzle is evident after 1970 (each sub-period coefficients is significantly different for the one of the preceding sub-period). These first results would seem to suggest that the most recent developments in the gravity literature on both theoretical and econometric sides — controlling for the zeros trade flows or including multilateral resistance terms —are unlikely to explain the puzzle observed since 1970.

But there are obvious limitations to results obtained from a meta-analysis since the estimates we do have not enough points to really appreciate the evolution of the distance effect *within* a sample, *within* a specification or *across* econometric methods (see Disdier and Head, 2008 for further discussion of other shortcomings).[14] We return to these issues in section 4 where we explore more systematically the evolution of $\tilde{\theta}$ in an integrated framework (i.e. *within* a sample, gravity model specification or econometric method).

[14] For instance, over the 1467 point estimates, only 52 (from 7 different studies) concerns developing countries.

Figure 2. Distance Puzzle in existing empirical literature

Figure 2a

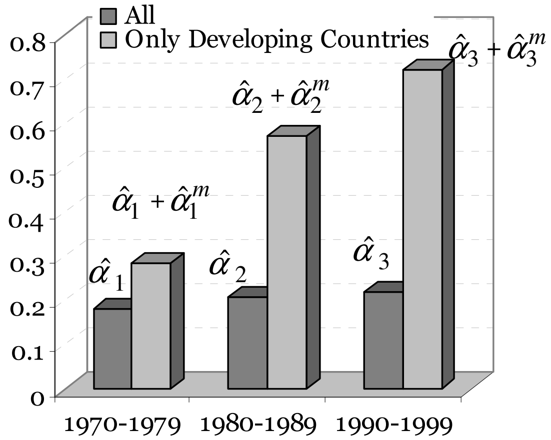


Figure 2b

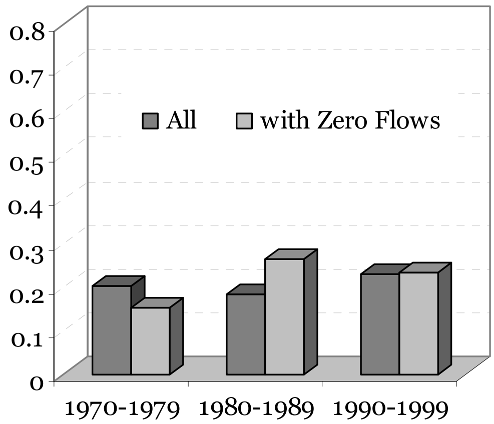
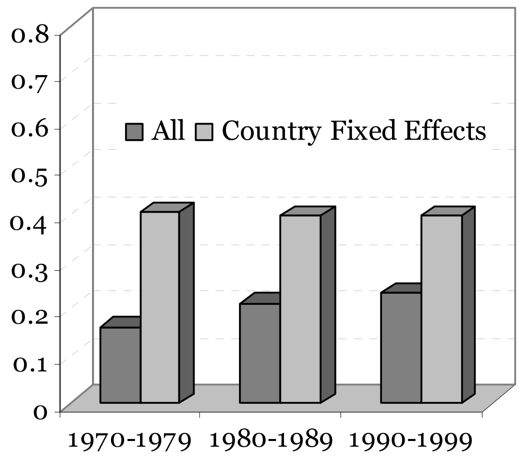


Figure 2c



Source: authors' calculations from the Disdier and Head (2008) database.

3. The Regionalization of Trade for Low-income countries

The gravity model with separability in trade costs has two important implications for the evolution of trade. First, it predicts that a relative fall in border-related costs should lead countries to increase the volume of international trade (relative to internal trade). This prediction is largely borne out by the data: since 1980, world production has increased by 75% while international trade has increased by 300% (Berthelon and Freund, 2008). Second, a reduction in all costs related to distance (including better information about distant markets) should lead countries to increase their volume of trade with distant partners, while on the contrary, if the relative costs associated with distance increase, countries should trade with closer partners. This implication of cost minimization was exploited by Carrère and Schiff (2005) who computed the average distance of trade (ADOT) directly from the bilateral trade data at successive points in time and more recently by Berthelon and Freund (2008) who computed a measure of potential trade ($ADOT^P$) predicted by relative country size. The measures are:

$$ADOT_t = \sum_i \sum_j \frac{X_{ijt}}{X_{wt}} D_{ij} \qquad (9)$$

where X_{ijt} are exports from *i* to *j* in t, X_{Wt} are world exports in t, and D_{ij} is distance between *i* and *j*. The corresponding potential measure is the gravity-predicted bilateral trade in a frictionless world where the volume of bilateral trade is proportional to the product of the countries GDPs (denoted $Y_{(i)t}$):

$$ADOT_t^P = \sum_i \sum_j \frac{X_{ijt}^p}{X_{wt}^p} D_{ij} \quad ; \quad X_{wt}^p = \sum_i \sum_j X_{ijt}^p = \sum_i \sum_j \frac{Y_{it} Y_{jt}}{Y_{wt}} \qquad (10)$$

This measure will change only as a result of changes in the dispersion of incomes around the world and it will be maximal if all countries have the same size. So, in a gravity world, a higher potential trade for a group of countries simply means less dispersion in economic size in that group. Feenstra (2004, chp. 5) reports results showing that this measure of potential trade fits the data quite well for developed countries but less for developing countries.

Then, if gravity is an adequate description of the volume of bilateral trade, the ratio of potential to actual trade is a measure of trade costs. Since it is suggested by the gravity model we call this ratio the average distance ratio (ADR):

$$ADR_t = ADOT_t^P / ADOT_t \qquad (11)$$

The values of these ratios are reported for our sample of 124 countries over the period 1970-2006.[15] To iron out fluctuations, each point is a 5-year

[15] The sample includes all countries except microstates and ex-FSU countries giving us a balanced sample (see the list of countries in appendix A1, table A1.1). But using the complete sample of 190 coutries does not change the results presented here and in the rest of the paper (results available upon request). Nominal bilateral trade flows (in US$,

average. Because of the preliminary results suggesting that low-income countries are different, we also report averages for the richest and poorest tercile of countries (each tercile has 39 observations). To ease the reading, we set ADR_t to 1 in for the period 1970-1974. It can be seen that the ADR ratio is quite stable fluctuating around the value of 1 for the whole sample, even though the small decline could be taken to suggest that barriers to trade have been increasing in relative terms, leading countries to shift trading patterns towards closer partners.

As suggested by the systematically higher estimates of the elasticity of trade to distance for developing countries in the meta-analysis, we also report separately the potential and actual distance of trade for the poorest third (39 countries) in the 124 in the sample with all countries in the sample. [16] By selecting all trade of the poorest countries rather than trade with only the remaining (85) countries, we are addressing directly the issue of whether their trade is becoming more regionalized, regardless of the partners.

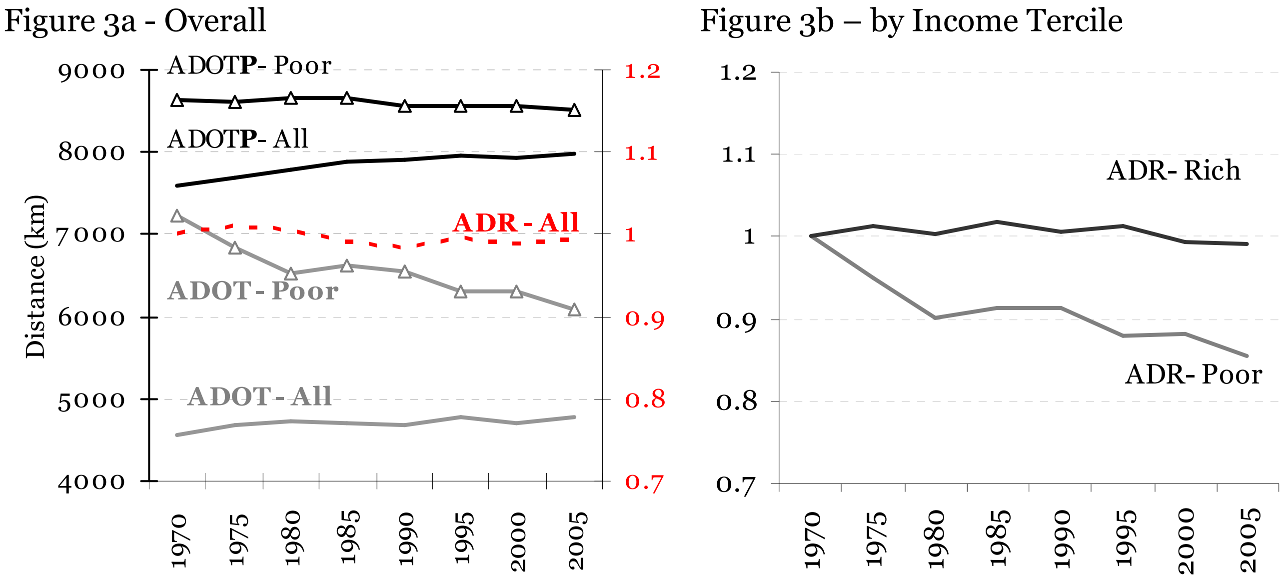
Figure 3 shows a higher potential trade for the poorest countries, reflecting larger dispersion in incomes for this group. More interesting is the large fall in the average distance of trade for this group implying that poor countries have shifted their trade towards geographically closer partners. For this group the average distance of imports fell by more than 15% from 7200 kms in 1970-1974 to 6000 in 2005-2006.

---

c.i.f), are taken from UN-COMTRADE (via WITS), divided by the US deflator. We use import data as it is well-recognized that they are more accurately reported by the customs authorities. For developing country, we use mirror estimates, i.e. export data reported by partner countries. GDP and population are taken from the World's Bank World Development Indicators 2008. Distance measures and dummy variables indicating whether the two countries are contiguous, share a common language, or have had a common colonizer are from the Centre d'Etudes et de Prospectives et d'Informations Internationales (CEPII). The simple distances are calculated following the great circle formula, which uses latitudes and longitudes of the most important city (in terms of population).

[16] The list of countries is given in appendix A1, table A1.1. The computation of terciles was carried out by splitting the country in three groups on the 1970-2006 average. We then checked if the classification would have changed if we had used beginning or end-of-period GDP figures. Concerning for instance the poorest tercile, compared to the list reported in table A1.1, China and Sri Lanka would have been included in this group at the beginning of the period (instead of Haiti and Zimbabwe) while Ivory Cost Only would have been included in this group (instead of Pakistan) at the end. Note that the poorest tercile matches perfectly the low-income country group as defined by the World Bank in 2006.

Figure 3. Average distance and Indirect Trade Cost Measures for 124 countries, 1970-2006



Note: 5-years periods over 1970-2004 and a 2-years period 2005-2006
Source: authors' calculations on data from UN-COMTRADE and WITS.

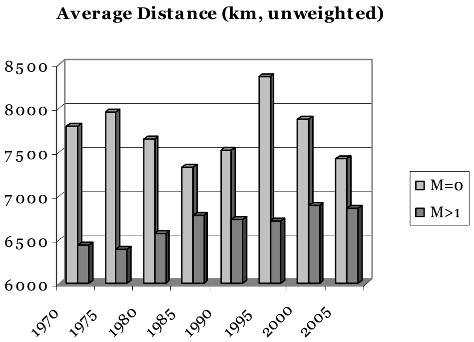
This changing pattern is reflected in diverging paths of the ADR ratios for the whole sample and for the lowest tercile. If the gravity model is an adequate of representation of the determinants of trade, then we get in figure 3a confirmation of the meta-analysis results of figure 2a: the costs of barriers to trade for the poorest countries have gone up in relative terms with a fall of 15% in the average distance of trade over the sample period.

Why did poor countries switch to closer trade partners? [17] In this simple setting, the only two possibilities are changing weights of existing trading partners, or changing trading partners. First, it could be that close trading partners (e.g. China and India in Asia) grew fast. This would result in the observed regionalization of trade for the poorest tercile. If so, we should then also observe a decrease in the average potential distance because of the increasing GDP weights for the close partners. However, figure 3a indicates that the potential distance of trade barely increases, so this effect cannot be a major factor.

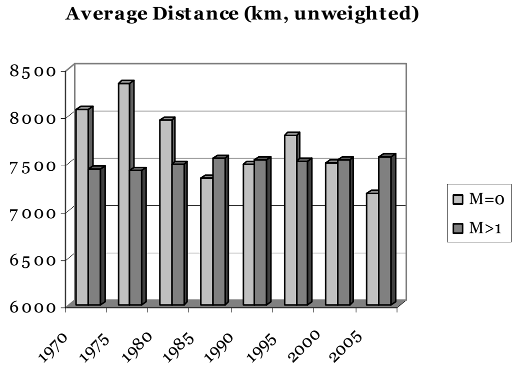
[17] To illustrate this point, we report in appendix A4 the 3 main import suppliers of each of the 39 poorest countries of the sample with their trade share and distance, for 1970-1975 and 2005-2006

Figure 4 Average distance of zero trade flows, Richest and Poorest Terciles.

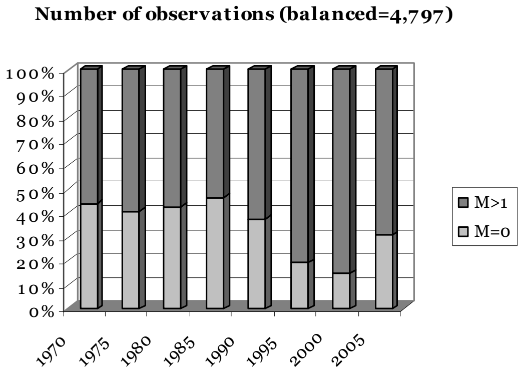
Poorest Tercile (39)



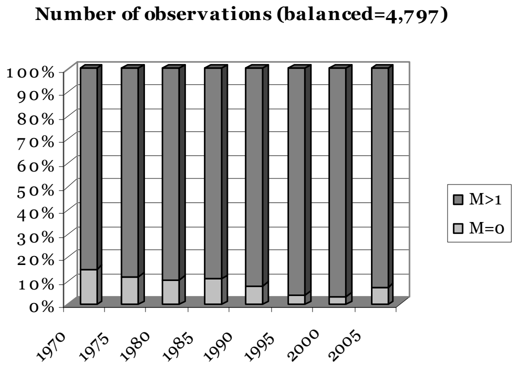
Richest Tercile (39)



Poorest Tercile (39)



Richest Tercile (39)



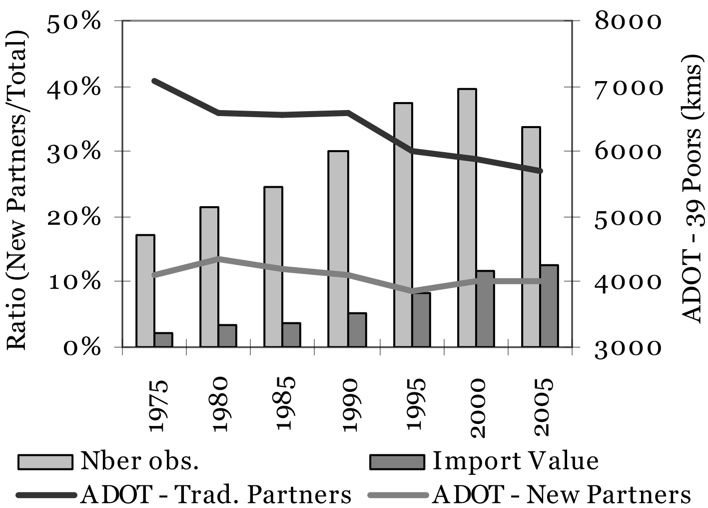
Note: 5-years periods over 1970-2004 and a 2-years period 2005-2006

Source: authors' calculations on data from UN-COMTRADE and WITS

The other possibility is a change in the composition of trading partners. Indeed, as shown in figure 4, over the sample period, the number of zero trade flows is quite stable until 1990, around 45% for the poorest tercile and 15% for the richest tercile. [18] Then the number of zero trade flows decreases sharply and fell by half. In contrast to the richest tercile, the average distance of zero trade flows for the poorest tercile is consistently higher than for positive trade flows, but the gaps narrow in later years.

The effect of this expansion of trade and its implication for the average distance of trade is shown in figure 5 for the poorest tercile. The figure disaggregates the ADOT for the lowest tercile into the two components: (i) "traditional", i.e. existing trade partners with already positive trade flows in 1970-1974 (intensive margin); (ii) "new" trade partners with positive trade flows since 1975 (extensive margin). We also report the weights of each trading partner (or each margin) in total trade.

Figure 5: Average Trade Distance of Poorest countries with Traditional and New Trade Partners



Two patterns are evident: first the regionalization of trade is partly reflecting the closer distance of the "new" partners that are significantly closer than the existing partners and they have an increasing weight in the

[18] We observe in the mid-1980s a slight increase in the number of zero trade flows compared to preceding years mainly for the poorest country' trade. This slowdown in trade growth around the mid-1980s is also visible in Felbermayr and Kohler (2006, figures 3a and 3b based on Rose's database) and in Helpman, Melitz and Rubinstein (2008, figures 1 and 2, based on Feenstra's database). As these new (and temporary) zero trade flows also concern geographically close partners, this results in a decrease of the unweighted average distance of zero trade flows during the 1980s.

total value of imports. Hence, part of the puzzle is along the extensive margin. Second, within the existing "traditional" group, the poorest countries have shifted towards geographically closer partners.[19] In the absence of data on the product composition of new trade partners, it is difficult to know what leads developing countries to choose closer trading partners since it could reflect a change in product composition as new products are initially shipped to close partners. In any case, it is clear from figure 5 that the regionalization of trade is also generated by trade redistribution within the intensive margin. [20]

The conclusion from this inspection of the raw data is that the poorest countries have shifted trade towards physically closer trading partners which would be expected from gravity theory if the relative trade costs with physically closer partners fell more than trade costs with further-away partners. This could be the case if the closer partners are those who reduced most their barriers to trade. In addition, even though on average partners with zero trade are further away than partners with positive trade, when extending trade to new partners, the poorest countries have selected those countries that are closest. Both patterns are consistent with a minimization of trade costs in a formulation in which distance matters. These patterns could also have resulted from the proliferation of regional trade agreements among the poorer countries.

The next section explores if this increasing elasticity of trade to distance only applies for the poorer countries in the sample after controlling for some of the factors that could alter distance-sensitive trade costs.

4. THE PERSISTENT RISING DISTANCE EFFECT FOR LOW-INCOME COUNTRIES

If the gravity model is an adequate representation of bilateral trade, one should obtain increasing values of the elasticity of trade to distance (the "distance effect"), $\tilde{\theta}$, over time in the gravity model, but only for developing countries. This is indeed what comes out of the alternative estimates below: repeated cross-sections (each cross-section representing a 5 years average) as in the vast majority of cases reported in the meta-

[19] We checked that the patterns described are robust when we drop alternatively India (included as reporter in the group of the 39 poorest countries and dominates other countries in terms of trade value) and China (included in the 123 partner countries) from the sample. Results of figure 5 are unchanged (available upon request).
[20] In a cross-section estimation of trade costs on the number of products exported by a country, Allen and Shepherd (2007) find robust and quantitatively significant positive estimates of the number of product varieties to a battery of trade costs with lower trade costs leading to increases in product varieties.

analysis; and panel estimates in which the estimated coefficient of distance is allowed to vary over time. The panel formulation is more suitable to incorporate time-dependent trade costs identified in (5) which we do when we build a time-series index of the quality of infrastructure. The panel estimates also allow for a better control for omitted variables by using country-pair specific effects. Hence, under panel estimation, omitted bilateral effects are no longer captured by the distance coefficient. On the other hand, because the number of zero trade flows is important for most of the sample period (especially for the poorest tercile), it is useful to explore several methods for controlling for zero values. This is better done in cross-section than in panel. We start with cross-section estimates, then turn to the panel estimates.

4.1 Cross-section Estimates

We follow the by-now standard approach and estimate:

$$\ln\left(M_{ij}\right)=\alpha_0+\alpha_i+\alpha_j+\theta\ln\left(D_{ij}\right)+\sum_m\lambda_m\ln\left(z_{ij}^m\right)+\nu_{ij} \qquad (12)$$

where α_i and α_j are the importer and exporter fixed effects that capture the multilateral resistance variables, and all other variables that are country specific and that will appear in the panel estimates: GDPs', multilateral term indexes and indices of the quality of infrastructure.[21]

Results are reported in table 1 for the first and last periods under different estimation methods to account for the zero trade flows in the data with the evolution of the estimated distance elasticity reported in figures 6 and 7.[22] Column (1) serves as a reference and reports OLS estimates in which the zero trade flows are considered as missing variables (this corresponds to the majority of estimates reported in the meta-analysis). Column (2) reports the results from the standard solution in the literature using $\ln\left(1+M_{ij}\right)$ instead of $\ln\left(M_{ij}\right)$ as the dependent variable (see e.g. Frankel, 1997). This increases the mean value of exports by one unit without affecting its variance and, with this correction, country-pair with zero trade flows are represented by a zero value of the dependent variable ($\ln\left(1+M_{ij}\right)$). However, the OLS estimator does not take into account the

[21] In this sample with low-income countries it is preferable to use OLS rather than the systems approach used by AvW to avoid the measurement error associated with the multilateral resistance variables. The log-linear approximation proposed by Baier and Bergstrand (2009) will be used in the panel estimates reported below.

[22] All results are reported in appendix A3, table A3.1.

censorship of the dependent variable. Column (3) reports the results from the Eaton and Tamura (ET, 1994) tobit with $\ln(a+M_{ij})$ as dependant variable. Under the ET estimator, instead of arbitrarily imposing $a=1$, the value of the a parameter is endogenously determined and the dependent variable will be censored at the value $\ln(a)$.[23] Finally column (4) reports the results from the Poisson Pseudo Maximum Likelihood (PPML) estimator proposed by Santos Silva and Tenreyro (2006) to deal with heteroskedasticity in log-linear gravity models.[24] Because of the controversy about the way to deal with zeros in the data, it is useful to compare the estimates under the two estimators.[25] Note however, that if we control for the censorship bias due to the relatively large number of zeros in the data, we do not decompose the distance effect into within-intensive and extensive margins as proposed by Helpman, Melitz and Rubinstein (2008). While extending the estimation to explore this issue is interesting, and was already explored partly using the descriptive statistics in preceding section, a credible improvement would require a plausible identification variable for the first step for this sample over 1970-2006.[26]

---

[23] We also ran a Tobit estimator on $\ln(1+M_{ij})$ which produced similar results to those obtained with the Eaton-Tamura estimator (see appendix A3, table A3.1). For the (ET) Tobit estimator with country dummies, we did not use the complicated transformation for a fixed effects Tobit developed by Honoré (1992) but a "simple" pooled Tobit as developed in Wooldridge (2002, pages 540-542). See also Arellano and Honoré (2001, section 7).

[24] We fit conditional fixed effects PPML (see details in Arellano and Honoré, 2001, section 5).

[25] The PPML estimator handles the inconsistency introduced by log-linearization model in the presence of heteroskedasticity, but it is not clear that it is a better estimator to handle the presence of zeroes in trade data. Using Monte-Carlo simulations, Santos Silva and Tenreyro show that that the PPML estimated of (12) with M_{ij} as dependent variable produce estimates with the lowest bias for different patterns of heteroskedasticity (compared to OLS on $\ln(1+M_{ij})$ and a tobit on $\ln(a+M_{ij})$, see their table 5 page 30). However, Martin and Pham (2008) pointed out that Silva and Tenreyro used a data-generating-process for their Monte Carlo analysis which is a fundamentally different data generating process from that underlying the zero values in models of Eaton and Tamura (1994) or Helpman, Melitz and Rubinstein (2008). When correcting the data-generating process, Martin and Pham (2008) find that the Eaton and Tamura (1994) Tobit estimates have a lower bias than those obtained with the PPML estimator.

[26] To do so would require a two-stage equation estimation procedure with a selection equation for the decision to trade across partners (identifying the extensive margin) followed by a trade flow equation in the second stage (identifying the intensive margin). Helpman, Melitz and Rubinstein (2008) use "regulation costs of firm entry" as identification variable for their 1986 cross-section regression. However, yearly (or five-year average data) necessary for a credible identification strategy, are not available.

Table 1. Barriers to Trade: Cross-section results

Methods	OLS	OLS	ET- Tobit	PPML	OLS	OLS	ET- Tobit	PPML
dependent var.	ln(M)	ln(1+M)	ln(a+M)	M	ln(M)	ln(1+M)	ln(a+M)	M
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			1970				**2005**	
lnDij	-1.066***	-1.115***	-1.231***	-0.668***	-1.429***	-1.336***	-1.325***	-0.710***
	(0.0352)	*(0.034)*	*(0.034)*	*(0.000)*	*(0.0303)*	*(0.038)*	*(0.029)*	*(0.000)*
Common Border	0.898***	0.326*	0.244*	0.367***	0.821***	0.613***	0.584***	0.555***
	(0.142)	*(0.178)*	*(0.132)*	*(0.000)*	*(0.146)*	*(0.188)*	*(0.115)*	*(0.000)*
Common Language	0.702***	1.001***	1.201***	0.328***	0.967***	1.112***	0.967***	0.203***
	(0.0638)	*(0.058)*	*(0.063)*	*(0.000)*	*(0.0585)*	*(0.064)*	*(0.053)*	*(0.000)*
Colonial links	1.315***	1.401***	1.329***	0.589***	0.639***	0.663***	0.695***	0.00188***
	(0.114)	*(0.134)*	*(0.161)*	*(0.000)*	*(0.109)*	*(0.142)*	*(0.141)*	*(0.000)*
Nber Obs.	10,403	15,252	15,252	15,252	13,384	15,252	15,252	15,252
% of zero Trade flows	0%	32%	32%	32%	0%	12%	12%	12%
R^2 or pseudo-R^2	0.698	0.745	0.743	0.877	0.799	0.817	0.910	0.905

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.
Fixed country effects are not reported.
Source: authors' calculations

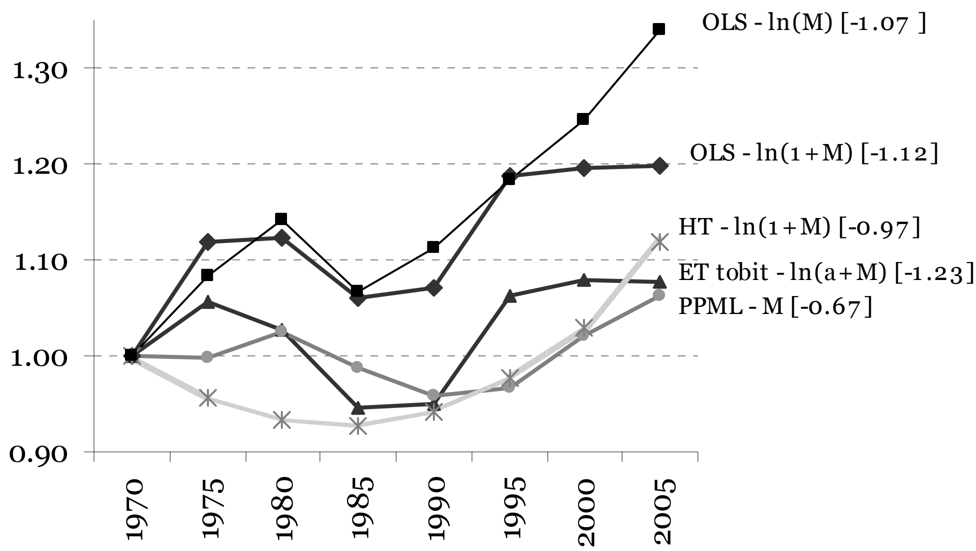
Several patterns stand out in table 1. First, as expected, the dummies have the usual signs and usual significance levels. There are however some changes in the coefficient values over the sample period reflecting expectations from a globalizing world. The value of the common language coefficient falls drastically through time: based on columns (4) and (8) (the PPML estimation), sharing a common language increases trade by 39% in 1970-74 but only by 22% in 2005-06. Colonial links also become far less important quantitatively over the period, increasing trade by around 80% in 1970-74 but only 0.2% in 2005-06.

The estimates of θ are high, but well in the range of values reported in figure 1. Importantly, the PPML elasticity estimate is much lower and more plausible than the values obtained with the other estimators. Santos Silva and Tenreyro (2006) explain this systematic difference in the estimated value between the OLS and the PPML estimators by the heteroskedasticity rather than by the censorship bias. [27] The explanatory power of all the models reported in Table 1 is quite high and increase over time, suggesting that the gravity model is a better representation of bilateral trade in later years.[28] One reason for this better fit would be

[27] See the discussion in Santos Silva and Tenreyro (2006). Since the coefficient on the PPML represents the marginal estimate of a change in distance on bilateral trade, the coefficients also represent trade elasticities to distance. Note that the OLS on ln(1+M)) and ET tobit estimates give very close results on this sample.
[28] R^2 are not directly comparable to pseudo-R^2 as the number of observations is not always the same and more importantly, the R^2 are based on sums of squares whereas the peudo-R^2 are based on ratios of log-likelihoods.

better data, especially for developing countries. Another, would be a change in the trade structure of developing countries as development leads to a shift towards trade in differentiated rather than homogenous products, hence towards a situation closer to that depicted by the gravity model (see e.g. the evidence in Feenstra, 2004 and Evenett and Keller, 2002).

Figure 6. Cross-Section Estimates of the Elasticity of Trade to Distance θ



Note: 5-years periods over 1970-2004 and a 2-years period 2005-2006
Corresponding Trade elasticity to distance in 1970 reported into brackets.
Source: authors' calculations

Finally, the time-plot of the estimates of θ in figure 6 (with all distance coefficients normalized to unity on the first sub-period 1970-1974) confirms the existence of a puzzle. The puzzle is robust across estimators. Two conclusions come out of this comparison. First, the strong distance puzzle obtained in literature is partly due to the fact that, until recently at least, zero values were not handled by OLS estimates with no specific correction for the censorship of the sample. However, even after controlling for the zero trade flows, the distance puzzle remains highly significant. Second, the range of estimates obtained across the different methods produces a rather narrow range of estimates with the burden of distance on trade significantly higher at the end of the period in the range [+6.3%; +7.6%]. Taken together, these results confirm that the distance puzzle holds up to the scrutiny of typical econometric problems.

To check whether the increasing values of θ is attributable to the presence of developing countries, we re-estimate the (12) by introducing dummy variables for the richest and poorest terciles, i.e. we regress:

$$\ln\left(M_{ij}\right)=\alpha_0+\alpha_i+\alpha_j+\theta\ln\left(D_{ij}\right)+\theta^I\ln\left(D_{ij}\right).I_{ij}+\sum_m\lambda_m\ln\left(z_{ij}^m\right)+v_{ij} \quad (13)$$

with alternatively the dummy I_{ij} equals to 1 if:

- i *or* j belongs to the third poorest countries in the sample;
- i *and* j belongs to the third richest countries.

Figure 7. Cross-Section Estimates of θ by Tercile

Figure 7a. 39 poorest' trade with all countries

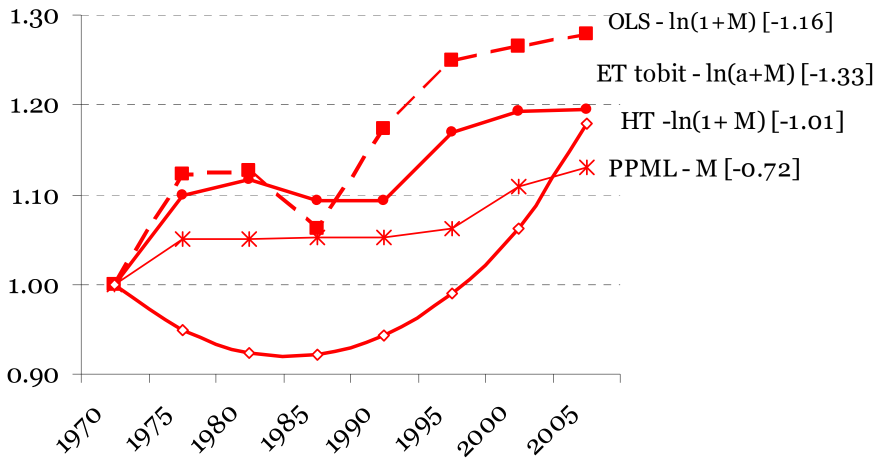
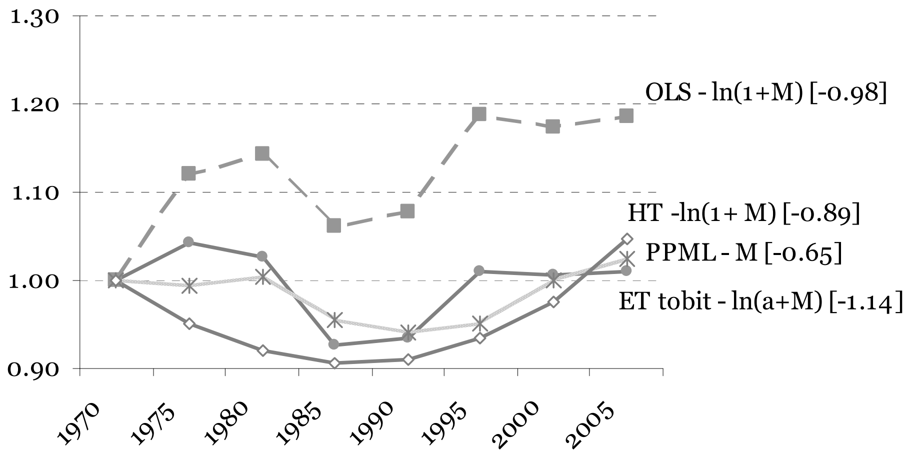


Figure 7b. Within 39 richest countries



Note: 5-years periods over 1970-2004 and a 2-years period 2005-2006, Corresponding Trade elasticity to distance in 1970 reported into brackets.
Source: authors' calculations

The results of this estimation are displayed in figures 7. It is clear that the results are due to the presence of developing countries in the sample since the estimates of θ in figure 7b show no more significant distance effect increase for the richest tercile and this is robust across estimators (except for the OLS estimates which are biased). By contrast the estimate of θ increases in the range [+13%; +19%] for the poorest tercile. Taken at face value, these estimates suggest that a doubling of distance would reduce poorest country's trade by 60% in 1970-74 and by 67% in 2005-2006 according to the ET tobit results (and respectively 39% and 43% in the PPML regressions).

4.2. Panel Estimates

Panel estimates allow the introduction of trade costs that vary through time and that are lumped together in (5) under the term τ_{ij} assumed to capture the ad-valorem equivalent of trade costs such as tariffs, communication and other transaction costs. In the absence of time-series data on border measures and other transaction costs, we follow earlier contributions (e.g. Limao and Venables , 2001 and Brun et al., 2005) and construct a time-series index of the quality of infrastructure for each country which becomes the new proxy for the bilateral transport costs, t_{ijt}. The augmented distance-dependent trade cost function becomes:[29]

$$t_{ijt} = \left(D_{ij}\right)^{\rho_1} \left(K_{it}\right)^{\rho_2} \left(K_{jt}\right)^{\rho_3} \qquad (14)$$

As constructed (see appendix A2), larger values for $K_{i(j)t}$ indicate a better infrastructure. [30]Again, the choice of functional form matters. If the cost function was additive with the infrastructure component independent of distance, the elasticity of transport costs to distance could increase if the fixed cost component were falling sufficiently faster than the variable component.[31]

[29] Since bilateral specific effects capture the time-invariant characteristics of bilateral trade, the trade cost function no longer includes the colonial, language and border dummy variables. Nor does it include fuel costs due to the year dummies.

[30] To compute the infrastructure index, we use data from the telecommunication sector (number of main telephone lines per 1000 workers), and the transportation sector (the length of the road and railway network —in km. per sq. km. of land area) and an index of the quality in the service of transport (the share of paved roads in total roads) from Canning (1998) and World Development Indicators Database (see details in appendix A2).

[31] See Brun et al. (2005, appendix D).

This trade cost function is introduced in (3) and (4) with country-pair fixed effects (FE) to capture the time-invariant characteristics of bilateral trade. Since these effects were not captured by the FE in the cross-section estimates, the panel estimates offer another robustness check on the earlier results although this is at the cost of imposing a trend specification for the evolution of distance. In particular, the country-pair fixed effects control for the North-South differences that Melitz (2007) found to reduce the estimate of θ by half over the period 1970-1995.[32] Furthermore, the panel specification allows for some asymmetry in the gravity equation since the bilateral specific effect on imports of *i* from *j* can be different from the corresponding bilateral effect of imports of *j* from *i* (see equation below).[33] Finally, year effects are included to capture all year shocks common to all country pairs such as variations in the cost of fuel (arguably a main factor affecting the marginal cost of transport).

To account for the variation of the multilateral resistance terms through time, we adopt the linear approximation proposed by Baier and Bergstrand (2009) to obtain unbiased and consistent reduced-form estimates (see the details in appendix A5). The estimated equation is:

$$\begin{aligned}\ln\left(M_{ijt}\right) = \lambda_t + \lambda_{ij} + \beta_1 \ln\left(Y_{it}\right) + \beta_2 \ln\left(Y_{jt}\right) - \theta \ln\left(D_{ij}\right) + \beta_3 \ln\left(K_{it}\right) + \beta_4 \ln\left(K_{jt}\right) \\ + \beta_5 \ln\left(MR_{it}\right) + \beta_6 \ln\left(MR_{jt}\right) + \varepsilon_{ijt}\end{aligned} \quad (15)$$

where

$MR_{it} = \left[\sum_k \frac{Y_{kt}}{Y_{wt}} \ln(D_{ik})\right], MR_{jt} = \left[\sum_k \frac{Y_{kt}}{Y_{wt}} \ln(D_{kj})\right]$ are the multilateral resistance or "remoteness" variables,[34] and λ_t is a vector of year dummies, λ_{ij} the bilateral fixed effects (with $\lambda_{ij} \neq \lambda_{ji}$) and ε_{ijt} the error term.[35] A quadratic time trend:

---

[32] Melitz (2007) regresses in OLS an equation very close to our OLS on ln(M) on 158 countries over 1970-1995 (five-years sub-periods). He obtained an increase in distance coefficient of around +18.8% while we find, over the same period +18.3% (using a random effect Hausman-Taylor specification on ln(M), see below). Once he controlled for North-South distance, Melitz obtained an increase in distance of around +8%, while with our corresponding random effect HT specification on ln(M) over 1970-1995 yields an estimated increase of +4%.

[33] Helpman, Melitz and Rubinstein (2008) motivate their model allowing for zero and asymmetric trade flows on data indicating asymmetry in trade flows over 1970-1997.

[34] Virtually the same results are obtained with the ad-hoc remoteness variable which uses the logs of the GDP weighted average distance rather than the GDP weighted average of the logs of the bilateral distance.

[35] Note that the random effect specification implies $\varepsilon_{ijt} = \mu_{ij} + \upsilon_{ijt}$ with μ_{ij} is a specific bilateral random effect, and υ_{ijt} is the idiosyncratic error term with the usual properties.

$$\theta \equiv \left| \left(\frac{\partial M_{ijt}}{M_{ijt}} \middle/ \frac{\partial D_{ij}}{D_{ij}} \right) \right| = \theta_1 + \theta_2.t + \theta_3.t^2 \tag{16}$$

is introduced in(15) to detect any significant evolution of $\hat{\theta}$. Because estimation in panel with bilateral fixed effects (FE), we also use a bilateral random effects (RE) estimator to obtain an estimate of the impact of distance on trade. Because the GDPs ($Y_{i(j)t}$) and infrastructure indices ($K_{i(j)t}$) are correlated with the bilateral random effects (μ_{ij}), we correct for this endogeneity using the Hausman-Taylor (1981) instrumental variables method.[36]

Results are reported in table 2. Note first that the FE and RE estimates in cols. 1 and 2 are very close indicating that endogeneity issues have been handled adequately (this is the case for all the variants). Signs and magnitude of coefficients are plausible. As suggested by the theory, the elasticity of trade with respect to income is significant and close to unity. The negative coefficients for the population variables confirm that larger countries are more self-sufficient (or that poorer countries trade less). It could also reflect that for large countries, the costs of trading with themselves rather than with others is relatively less. The multilateral resistance variables also have the expected positive sign. Thus given the absolute distance between i and j, the further country i is far from its trade partners, the more country i will trade with j. As expected, an improvement in the quality of infrastructure increases significantly the volume of trade. One could also interpret these positive coefficients in the broader sense of proxies for the quality of social infrastructure (physical infrastructure is largely a public good that will be underprovided in countries with poor social infrastructure). The distance coefficient in the HT specification is close to unity when taking into account the zero trade flows ($\ln(1+M_t)$).

The estimated trend from column (4) is reported in figure 6. It is slightly higher than the trend estimated with the ET tobit or the PPML (+11% vs. 7.6% and 6.3% respectively). Using the same approach as in (13), we also estimate the trend for the richest and poorest terciles and report the results in figure 7. Again, the results are quite close to those obtained by the repeated cross-section estimations. While it could be argued that not accounting for the censorship of zeroes in the panel estimates might make

[36] We carry out a pretest estimator based on Hausman (1978) (see Baltagi, Bresson and Pirotte (2003) to choose between the fixed effects, random effects, and HT estimators. The test statistics lead us to choose the HT estimator. See Brun et al. (2005) appendix B for the construction of instruments and the choice of estimator.

a difference, given the close values across estimators in the cross-section results in table 1, it is unlikely that not accounting for truncation would have significantly changed the results here.

Table 2. Panel estimations

Methods		FE	HT	FE	HT
dependent var.		ln(1+Mijt)	ln(1+Mijt)	ln(1+Mijt)	ln(1+Mijt)
		(1)	(2)	(3)	(4)
GDP i	*lnYit*	0.754***	0.763***	0.818***	0.813***
		(0.009)	*(0.009)*	*(0.031)*	*(0.008)*
GDP j	*lnYjt*	0.906***	0.914***	0.959***	1.011***
		(0.009)	*(0.009)*	-	*(0.008)*
Distance	*lnDij*	-	-0.989***	-	-1.071***
		0.000	*(0.024)*	*0.000*	*(0.023)*
Pop i	*lnPOPit*			-0.590***	-0.352***
				(0.052)	*(0.010)*
Pop j	*lnPOPjt*			-0.212***	-0.169***
				(0.052)	*(0.010)*
Multil. Resistance i	*lnMRit*	2.075***	1.609***	1.567***	1.487***
		(0.083)	*(0.051)*	*(0.883)*	*(0.074)*
Multil. Resistance j	*lnMRjt*	1.573***	1.209***	1.135	1.230***
		(0.083)	*(0.051)*	*(0.066)*	*(0.074)*
Infra i	*lnKit*	0.208***	0.191***	0.268***	0.256***
		(0.010)	*(0.010)*	*(0.029)*	*(0.009)*
Infra j	*lnKjt*	0.176***	0.159***	0.281***	0.227***
		(0.010)	*(0.010)*	*(0.033)*	*(0.009)*
trend in Dist	*tlnDij*	0.00961***	0.0111***	0.0110***	0.0110***
		(0.001)	*(0.001)*	*(0.002)*	*(0.001)*
trend² in Dist	*t²lnDij*	-0.000376***	-0.000377***	-0.000382***	-0.000381***
		(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
Bilateral specific effects		*yes*	*yes*	*yes*	*yes*
Year dummies		*yes*	*yes*	*yes*	*yes*
N. obs		417,110	417,110	417,110	417,110
N. country pairs		12,432	12,432	12,432	12,432

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We carried out two further robustness tests on the trend estimates (not reported here). First, we checked the sensitivity of estimates to imposing unitary elasticities on the GDPs. The results are unchanged. Second, we controlled for composition effects that preoccupied Melitz (2007) using the share of primary exports to control for the composition of exports (comparative-advantage based products vs. differentiated products).[37] As expected, a larger share of primary exports is associated with lower bilateral trade confirming the stylized fact that freight rates are higher for primary products than for manufactures because of the weight factor. But the distance estimates remain unchanged.

In sum, the robustness checks from the panel results which allow for bilateral (and asymmetrical) specific effects do not alter the result that the elasticity of trade to distance has increased for the lower-income countries.

5. Conclusions

The increasing distance effect in bilateral trade, now well-established in the literature on gravity-based trade (hence the "distance puzzle") has been addressed in several papers. Among the explanations put forth, the proper specification and estimation of the gravity equation ranks first with focus on the econometric methods (e.g. the treatment of zero trade flows, heteroskedasticity). Other explanations have focussed on omitted variable bias (e.g. country fixed effects or the introduction of multilateral resistance terms), and on sample selection (developed vs. developing countries). Using several approaches, we confirm that the distance puzzle is related to the sample: only the bottom third exhibit in a sample of 124 countries exhibit a rising distance effect, with the estimate of θ increasing in the range [+13%; +19%] over 1970-2006 regardless of the method used. Inspection of the data confirms that this result is not spurious as the "average distance of trade" has only fallen for this group of poorest countries, both at the extensive (new trade partners) and intensive (existing trade partners) margins. Hence, the last thirty five years have witnessed a regionalization of trade for the low-income countries.

Several possibilities could be explored. First, this regionalization of trade could simply reflect a more rapid reduction in policy-related trade costs (including border-related costs and trade restrictions) for this group relative to others over the period. If so, this regionalization of world trade would be predicted by the gravity model as most low-income countries are geographically close. This indirect evidence (since we do not have time-series data on the evolution of trade costs) would be good news as it would mean a deepening integration of this group of countries into the World Trading System. Another possibility is that the fixed component of trade costs (communication costs, increasing containerization) has fallen more rapidly than the variable component of trade costs. A less optimistic view

[37] Bertelon and Freund (2008) decompose the change in the distance coefficient into compositional and distance sensitivity effects. Ad-valorem freight rates for SITC 0-4 are about 50% higher than for SITC 5-9 (see Hummels (1999 table 1).

can be design if we assume that a growing part of world trade is generated by vertical specialization and just-in-time production over the period. Actually, in this case, trade costs can be viewed as a growing impediment in the supply-chain production and then, if low-income countries' trade costs (in particular distance-dependant, such as high markups in international shipping) remain high compared to other developing countries' trade costs, then, the observed trade regionalization can be interpreted as a marginalization of these countries.

References

Aidt, T. and M. Gassebner (2008), "Do Autocratic States Trade less?" Mimeo, University of Zurich.

Anderson, J. and D. Marcouiller (2002) "Insecurity and the Pattern of Trade: An Empirical Investigation", The Review of Economics and Statistics, 84,342-52.

Anderson, J. and E. Van Wincoop (2003) "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, 93, 170-92.

Anderson, J. and E. Van Wincoop (2004) "Trade Costs", *Journal of Economic Literature*, 42(3), 691-751.

Arellano, M. and B. E. Honoré (2001), "Panel Data Models. Some Recent Developments", Handbook of Econometrics, 5, Elsevier Science, Amsterdam.

Baier, S.L. 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.

Baltagi, B.H., Bresson, G. and A. Pirotte (2003), "Fixed Effects, Random Effects or Hausman-Taylor? A Pretest Estimator", *Economics Letters*, 79, 361-369.

Berthelon, M., and C. Freund (2008), "On the Conservation of Distance in International Trade", *Journal of International Economics*, 75, 310-20

Brun J.-F, Carrère C., Guillaumont P. and De Melo J. 2005. "Has Distance Died? Evidence from a Panel Gravity Model" *World Bank Economic Review* 19(1): 99-120.

Buys, P., U. Diechman, and D. Wheeler (2006), Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa", WBWP#4097.

Canning, D. 1998. "A Database of World Infrastructure Stocks, 1950–1995." *World Bank Economic Review* 12(3):529–47.

Carrère and Schiff (2005) "On the Geography of Trade: Distance is Alive and Well", Revue Economique, 56, 1249-74.

Coe, D. Subramanian, A., and N. Tamisara (2007), "The Missing Globalization Puzzle", *IMF Staff Papers*, 54, 34-58.

Deardorff, A. (2003), "Time and Trade: The Role of Time in Determining the Structure and Effects of International Trade, with an Application to Japan," in Robert M. Stern, ed., Japan's Economic Recovery: Commerdial Policy, Monetary Policy, and Corporate Governance, MA: Edward Elgar Publishing Inc.,63-76.

Allen, D. and B. Shepherd (2007) Barriers to Entry, Trade Costs, and Export Diversification in developing Countries", Mimeo, World Bank.

Disdier A.C. and Head, K. 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *The Review of Economics and Statistics*.

Eaton, J. and A. Tamura, (1994), "Bilateralism and regionalism in Japanese and U.S. trade and direct foreign investment patterns." *Journal of Japanese and International economies,* 8, 478-510.

Estevadeordal, Antoni, Brian Franz, and Alan M. Taylor (2003), "The Rise and Fall of World Trade 1870-1939", *Quarterly Journal of Economics*, 68(2), 359-408.

Evenett, S. and Keller (2002), "On Theories Explaining the Success of the Gravity Equation", *Journal of Political Economy*, 110(2), 281-396.

Feenstra, R. (2004) *Advanced International Trade*, Princeton, Princeton University Press.

Felbermayr, G.J. and W Kohler (2006) "Exploring the Intensive and Extensive Margins of World Trade", *Review of World Economics*, 112, 642-74.

Francois, J., and Manchin, M., 2007. "Institutions, Infrastructure and Trade", WBWPS# 4152.

Frankel, J. (1997) Regional Trading Blocs in the World Economic System, Institute for International Economics, Washington, D.C.

Friedman, T. (2005), The World Is Flat: A Brief History of the Twenty-First Century", eds. Farrar, Straus and Giroux.

Grossman G, (1998) "Comment on Deardorff" in J. A. Frankel ed. The Regionalization of the World Economy, Chicago, University of Chicago Press.

Hausman, A. and E. Taylor, (1981), "Panel data and unobservable individual effects," Econometrica, 49, 1377-1398.

Helpman, E., M. Melitz and Y. Rubinstein (2008) "Estimating Trade Flows: Trading Partners and Trading Volumes", *Quarterly Journal of Economics*, 123(2), 441-487

Honoré, B. E. (1992), "Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects", *Econometrica*, 60, 533-565.

Hummels, D. (1999), "Have International Transportation Costs Declined?" Mimeo, University of Chicago.

Hummels, D. (2001), 'Trade on Time", GTAP working paper 1152, Purdue University.

Hummels, D., V. Lugovskyy, A. Skiba (2009), "The Trade Reducing Effects of Market Power in International Shipping" *Journal of Development Economics,* forthcoming.

Leamer, Edward, and James Levinsohn (1995) "International Trade: The Evidence", in G.M. Grossman and K. Rogoff eds. Handbook of International Economics, vol. 3, New York, Elsevier.

Limaõ, N., and A. J. Venables. 2001. "Infrastructure, Geographical Disadvantage and Transport Costs." *World Bank Economic Review* 15(3):451–79.

Martin, W. and Pham, C. 2007. "Estimating the Gravity Model When Zero Trade Flows are Important", mimeo World Bank.

Melitz, J. (2007) "North, South and Distance in the Gravity Model", *European Economic Review*, 51, 971-991.

Melitz, J. (2008) "Language and Foreign Trade", *European Economic Review*, 52, 667-699.

Santos Silva, J.M.C and S. Tenreyro .2006. "The Log of Gravity". *Review of Economics and Statistics,* 88, 641-58.

Shepherd, John, and J. Wilson (2006) "Road Infrastructure Network in Europe and Central Asia: Does Network Quality Affect Trade?", WBWPS#4101.

Shepherd, John, and J. Wilson (2008) "Trade Facilitation in Asean Member Countries: Measuring Progress and Assessing Priorities", WBWPS#4615.

Wooldridge, J.M. (2002), Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press.

Appendices to
THE DISTANCE EFFECT AND THE REGIONALIZATION OF THE TRADE OF LOW-INCOME COUNTRIES

Céline Carrère
CERDI – CNRS

Jaime de Melo
University of Geneva and CERDI

John Wilson
World Bank

February 2009

Appendix A.1. Countries in the Sample

Table A1.1. Sample of 124 countries (i.e. without microstates and ex-FSU)

First Tercile (GDPpc<594 $)	Second Tercile (594$<GDPpc<4,163$)	Third Tercile (GDPpc>4,163$)	GDPpc missing
Bangladesh	Albania	Argentina	Afghanistan
Benin	Algeria	Australia	Cuba
Burkina Faso	Angola	Austria	Korea, Dem. Rep.
Burundi	Bolivia	Belgium-Luxembourg	Myanmar
Cambodia	Brazil	Canada	Somalia
Central African Republic	Bulgaria	Czech Republic	
Chad	Cameroon	Denmark	
Congo, Dem. Rep.	Chile	Finland	
Ethiopia(includes Eritrea)	China	France	
Gambia, The	Colombia	Gabon	
Ghana	Congo, Rep.	Germany	
Guinea	Costa Rica	Greece	
Guinea-Bissau	Cote d'Ivoire	Hong Kong, China	
Haiti	Dominican Republic	Ireland	
India	Ecuador	Israel	
Kenya	Egypt, Arab Rep.	Italy	
Lao PDR	El Salvador	Japan	
Liberia	Guatemala	Korea, Rep.	
Madagascar	Honduras	Kuwait	
Malawi	Hungary	Lebanon	
Mali	Indonesia	Libya	
Mauritania	Iran, Islamic Rep.	Mexico	
Mongolia	Iraq	Netherlands	
Mozambique	Jamaica	New Zealand	
Nepal	Jordan	Norway	
Niger	Malaysia	Oman	
Nigeria	Mauritius	Poland	
Pakistan	Morocco	Portugal	
Rwanda	Nicaragua	Saudi Arabia	
Senegal	Panama	Singapore	
Sierra Leone	Papua New Guinea	Spain	
Sudan	Paraguay	Sweden	
Tanzania	Peru	Switzerland	
Togo	Philippines	Trinidad and Tobago	
Uganda	Romania	United Arab Emirates	
Vietnam	South Africa	United Kingdom	
Yemen	Sri Lanka	United States	
Zambia	Syrian Arab Republic	Uruguay	
Zimbabwe	Thailand	Venezuela	
	Tunisia		
	Turkey		

Notes:
GDP per capita: average over 1970-2006, in constant 2005 $

Appendix A2. The Infrastructure Index

To compute the infrastructure index, we use a large panel data set comprising 144 countries, spanning the years 1962-2006. We use data from the telecommunication sector (number of main telephone lines per 1000 workers), and the transportation sector (the length of the road and railway network —in km. per sq. km. of land area) and an index of the quality in the service of transport (the share of paved roads in total roads) from Canning (1998) and World Development Indicators Database.

Reconciliation of the infrastructure series

We use 4 indicators
- the number of main telephone lines per 1000 workers;
- the length of the road network —in km. per sq. km. of land area;
- the length of the railway network —in km. per sq. km. of land area;
- the length of the road network —in km. per sq. km. of land area;

For these indicators, we use two database: (i) Canning (1998) covering 144 countries over 1960-1995 and (ii) the World Development Indicators Database with data essentially from 1980 to 2006. Table A2.1 reports, for each of the 4 indicators, the number of non missing observations in the Canning Database, in the WDI database and the number of overlapping observations (country / year observations that are non missing in both database).

To reconciliate the Canning and WDI database we proceed, for each of the 4 indicators and each country, in three steps:

(i) We compute the trend in the infrastructure indicator based on Canning database over 1962-1995 and, based on this estimated trend, we generate the predicted infrastructure for 1995-2006.

(ii) We compute the ratio of the WDI data on the corresponding predicted (or observed when overlapping) indicator based on the Canning database; we reconciliate the series if the ratio is in the range [0.8; 1.2], which is the case for 92% of the WDI observations, we put a missing value otherwise;

(iii) We reestimate a trend on the reconciliated data and complete the missing value with the predicted indicator.

Table A2.1 reports the proportion, in the reconciliated data, of observation from Canning, WDI and computed from the estimated trend.

Table A2.1: Number of non missing observations

		Rail		Road		Telephone		Paved Road	
Raw Data		(1)		(2)		(3)		(4)	
	Canning	4,288		3,254		4,141		3,283	
	WDI	1446		1494		3,642		310	
	Overlapping	875		402		2,435		67	
Final Data									
	Canning	4,288	66%	3,254	50%	4,141	64%	3,283	57%
	WDI	547	8%	997	15%	1,205	19%	198	3%
	Completed	1,645	25%	2,229	34%	1,134	18%	2,279	40%
		6,480	100%	**6,480**	100%	**6,480**	100%	**5,760**	100%

We have all information on Telephone, road and paved road for the whole sample over 1962-2006. Concerning the railway, data are missing over the whole period for 16 countries.[38]

[38] Bahamas, Belize, Bhutan, Canada, Germany, Lao, Portugal, Samoa, St Kitts and Nevis, St Lucia, St Vincent and Grenadine, Sudan, Tonga, United Arab Emirates, Uruguay, Vanuatu.

Computation

We choose to employ a single infrastructure indicator for the empirical analysis of bilateral trade for mainly 2 reasons. The first one is the high correlation among measures of various kinds of infrastructure -- telecommunications, road and railway networks. For example, we find that the correlation between measures of telephone and road density is 0.5, and between rail and railway density is close to 0.6. In a linear regression framework, this close association among different infrastructure categories makes it hard to obtain reliable estimates of the individual coefficients of variables representing different kinds of infrastructure assets. The second reason is that, due to the quality of the infrastructure data, it is certainly too ambitious to try to assess with accuracy the impact of each infrastructure components while using these data in an aggregate index allow us to capture the broad picture on country's infrastructure.

For these reasons, we build synthetic indices that summarize various dimensions of infrastructure and its quality. We propose 2 alternative indexes following:

(i) Limao and Venable (2001), Brun et al. (2005) and compute an simple average of the density of the road network, the paved road network, the rail network, and the number of telephones per person - we first normalize the variables to have the same mean, one, and then take the linear average over the four variables, ignoring the missing observations. This is equivalent to assuming that roads, paved roads, railways and telephones lines are perfect substitutes as inputs to a transport services production functions.

(ii) Calderon and Serven (2004) or Francois and Manchin (2007) and apply principal component analysis to disaggregate infrastructure indicators; our synthetic indices are given by the first principal component of the underlying variables. Note that before applying principal component analysis, the 4 underlying variables are standardized (mean 1 and variance 1) in order to abstract from units of measurement.

The first principal component of the 4 variables accounts for 67% of their overall variance and, as expected, it is highly correlated with each individual measure included. Specifically, the correlation between the first principal component and main telephone lines is 0.74, its correlation with the length of the railways network is 0.67, and with the length of the road and paved road network is 0.91 and 0.93 respectively. In addition, all 4 infrastructure variables enter the first principal component with approximately similar weights:[39]

$$K_{it}^{pca} = 0.5537 Road_{it} + 0.5661 PavedRoad_{it} + 0.4081 Rail_{it} + 0.4544 Tel_{it}$$

where K_{it}^{pca} is the synthetic index of infrastructure from the principal component analysis.[40]

Note that the correlation between the 2 infrastructure indexes (simple average and the first principal component) is 0.9971!

[39] The principal components are normed to 1 (sum of the squares of the weights is 1).

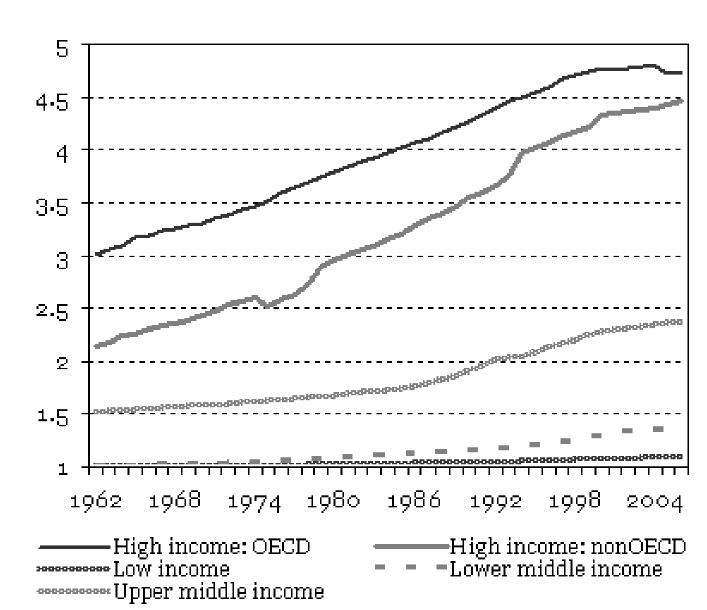
[40] For the 16 countries with missing data on the paved road network variables, we compute the first principal component of the 3 other variables:

$$K_{it}^{pca} = 0.5725 Road_{it} + 0.5598 Rail_{it} + 0.5990 Tel_{it}$$

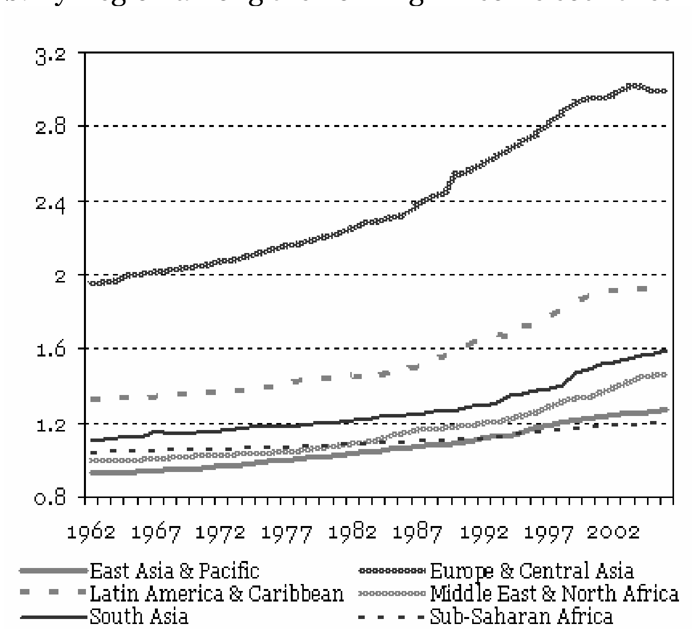
Descriptive statistics

Figure A2.1. Evolution of the infrastructure index, K_{it}^{pca} 1962-2006.

a. By Income



b. By Region among the non High Income countries



Appendix A3. Tables of Results

Table A3.1. Five-years periods – all sample (124 countries, i.e. 15,152 obs. per period)

Methods	OLS	Tobit	ET- Tobit	PPML	OLS	Tobit	ET- Tobit	PPML
dependent var.	ln(1+M)	ln(1+M)	ln(a+M)	M	ln(1+M)	ln(1+M)	ln(a+M)	M
	1970				1975			
lnDij	-1.115***	-1.380***	-1.231***	-0.668***	-1.148***	-1.394***	-1.249***	-0.666***
	(0.034)	*(0.044)*	*(0.034)*	*(0.000)*	*(0.036)*	*(0.044)*	*(0.032)*	*(0.000)*
Common Border	0.326*	0.224	0.244*	0.367***	0.345*	0.299	0.257**	0.329***
	(0.178)	*(0.217)*	*(0.132)*	*(0.000)*	*(0.176)*	*(0.210)*	*(0.127)*	*(0.000)*
Common Language	1.001***	1.479***	1.201***	0.328***	0.937***	1.289***	1.006***	0.259***
	(0.058)	*(0.075)*	*(0.063)*	*(0.000)*	*(0.060)*	*(0.074)*	*(0.060)*	*(0.000)*
Colonial links	1.401***	1.350***	1.329***	0.589***	1.317***	1.286***	1.260***	0.391***
	(0.134)	*(0.165)*	*(0.161)*	*(0.000)*	*(0.144)*	*(0.174)*	*(0.156)*	*(0.000)*
	1980				1985			
lnDij	-1.152***	-1.272***	-1.265***	-0.683***	-1.182***	-1.283***	-1.105***	-0.659***
	(0.038)	*(0.047)*	*(0.032)*	*(0.000)*	*(0.035)*	*(0.044)*	*(0.029)*	*(0.000)*
Common Border	0.151	0.0996	0.135	0.347***	0.141	0.111	0.188	0.502***
	(0.183)	*(0.222)*	*(0.125)*	*(0.000)*	*(0.173)*	*(0.215)*	*(0.115)*	*(0.000)*
Common Language	1.008***	1.361***	0.985***	0.232***	1.064***	1.425***	0.988***	0.320***
	(0.063)	*(0.079)*	*(0.059)*	*(0.000)*	*(0.060)*	*(0.078)*	*(0.054)*	*(0.000)*
Colonial links	1.165***	1.119***	1.117***	0.292***	1.030***	0.997***	0.980***	0.0669***
	(0.143)	*(0.177)*	*(0.153)*	*(0.000)*	*(0.134)*	*(0.172)*	*(0.137)*	*(0.000)*
	1990				1995			
lnDij	-1.195***	-1.311***	-1.169***	-0.640***	-1.325***	-1.403***	-1.309***	-0.646***
	(0.033)	*(0.039)*	*(0.027)*	*(0.000)*	*(0.031)*	*(0.033)*	*(0.026)*	*(0.000)*
Common Border	0.327**	0.328*	0.350***	0.639***	0.677***	0.655***	0.639***	0.719***
	(0.164)	*(0.193)*	*(0.109)*	*(0.000)*	*(0.156)*	*(0.166)*	*(0.104)*	*(0.000)*
Common Language	1.115***	1.367***	1.017***	0.309***	1.090***	1.205***	1.015***	0.227***
	(0.059)	*(0.071)*	*(0.051)*	*(0.000)*	*(0.055)*	*(0.060)*	*(0.048)*	*(0.000)*
Colonial links	0.945***	0.910***	0.919***	-0.00646***	0.727***	0.710***	0.737***	-0.0356***
	(0.124)	*(0.153)*	*(0.132)*	*(0.000)*	*(0.116)*	*(0.128)*	*(0.129)*	*(0.000)*
	2000				2005			
lnDij	-1.334***	-1.385***	-1.328***	-0.682***	-1.336***	-1.432***	-1.325***	-0.701***
	(0.033)	*(0.034)*	*(0.026)*	*(0.000)*	*(0.038)*	*(0.043)*	*(0.029)*	*(0.000)*
Common Border	0.776***	0.770***	0.740***	0.652***	0.613***	0.636***	0.584***	0.555***
	(0.156)	*(0.165)*	*(0.105)*	*(0.000)*	*(0.188)*	*(0.212)*	*(0.115)*	*(0.000)*
Common Language	1.106***	1.195***	1.042***	0.190***	1.112***	1.265***	0.967***	0.203***
	(0.054)	*(0.058)*	*(0.048)*	*(0.000)*	*(0.064)*	*(0.075)*	*(0.053)*	*(0.000)*
Colonial links	0.585***	0.566***	0.602***	-0.0135***	0.663***	0.668***	0.695***	-0.00188***
	(0.120)	*(0.128)*	*(0.131)*	*(0.000)*	*(0.142)*	*(0.164)*	*(0.141)*	*(0.000)*

Note:
5-years periods over 1970-2004 and 2-years subperiods 2005-2006.
importer and exporter dummies in all regressions.
Source: authors' calculations

Table A3.2. Panel estimations – all sample (190 and 124 countries)

Methods		FE	HT	FE	HT	FE	HT	FE	HT
dependent var.		ln(Mijt)	ln(Mijt)	ln(Mijt)	ln(Mijt)	ln(1+Mijt)	ln(1+Mijt)	ln(1+Mijt)	ln(1+Mijt)
sample		190 countries	190 countries	124 countries	124 countries	190 countries	190 countries	124 countries	124 countries
GDP i	*lnYit*	0.993***	0.988***	1.004***	1.021***	0.652***	0.630***	0.754***	0.763***
		(0.014)	*(0.013)*	*(0.015)*	*(0.014)*	*(0.007)*	*(0.006)*	*(0.009)*	*(0.009)*
GDP j	*lnYjt*	1.324***	1.329***	1.336***	1.367***	0.781***	0.759***	0.906***	0.914***
		(0.014)	*(0.012)*	*(0.015)*	*(0.014)*	*(0.007)*	*(0.006)*	*(0.009)*	*(0.009)*
Distance	*lnDij*	-	-1.417***	-	-1.530***	-	-0.650***	-	-0.989***
		0.000	*(0.038)*	*0.000*	*(0.040)*	*0.000*	*(0.020)*	*0.000*	*(0.024)*
Remoteness i	*Rit*	1.906***	1.554***	1.464***	1.166**	2.086***	1.846***	2.075***	1.609***
		(0.120)	*(0.080)*	*(0.127)*	*(0.084)*	*(0.067)*	*(0.042)*	*(0.083)*	*(0.051)*
Remoteness j	*Rjt*	0.388***	0.759***	0.605***	1.219***	1.901***	1.596***	1.573***	1.209***
		(0.118)	*(0.079)*	*(0.125)*	*(0.083)*	*(0.067)*	*(0.042)*	*(0.083)*	*(0.051)*
Infra i	*lnKit*	0.179***	0.163***	0.298***	0.280***	0.0531***	0.0539***	0.208***	0.191***
		(0.015)	*(0.014)*	*(0.017)*	*(0.016)*	*(0.007)*	*(0.007)*	*(0.010)*	*(0.010)*
Infra j	*lnKjt*	0.270***	0.256***	0.415***	0.406***	0.0157**	0.0160**	0.176***	0.159***
		(0.015)	*(0.015)*	*(0.017)*	*(0.016)*	*(0.007)*	*(0.007)*	*(0.010)*	*(0.010)*
trend in Dist	*t.lnDij*	-0.00288**	-0.00198*	-0.00181*	-0.00188*	0.00548***	0.00747***	0.00961***	0.0111***
		(0.001)	*(0.001)*	*(0.001)*	*(0.001)*	*(0.001)*	*(0.001)*	*(0.001)*	*(0.001)*
trend² in Dist	*t².lnDij*	-0.0000215	-0.0000362	-0.0000244	-0.0000309	-0.000321***	-0.000335***	-0.000376***	-0.000377***
		(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Bilateral specific effects		*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Year dummies		*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
N. obs		365,492	365,492	292,465	292,465	592,294	592,294	417,110	417,110
N. country pairs		16,974	16,974	12,039	12,039	18,360	18,360	12,432	12,432
R^2		0.138	.	0.159	.	0.151	.	0.18	.

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: authors' calculations

Appendix A4. Evolution of the 3 Main import suppliers of the 39 poorest countries of the sample

Reporter (Importer)		1970 -1974 Partner (Exporter)	Trade share	Distance (km)	2000 -2004 Partner (Exporter)	Trade share	Distance (km)	2005 -2006 Partner (Exporter)	Trade share	Distance (km)
Bangladesh	1	United States	27%	12680	India	16%	1422	India	20%	1422
Bangladesh	2	India	15%	1422	**China**	13%	3036	Singapore	10%	2888
Bangladesh	3	Japan	10%	4904	**Singapore**	10%	2888	Hong Kong, China	8%	2439
Benin	1	France	37%	4727	France	21%	4727	France	16%	4727
Benin	2	Germany	8%	5033	Thailand	6%	10704	Thailand	13%	10704
Benin	3	United States	7%	8417	**United Kingdom**	6%	5026	Côte d'Ivoire	9%	733
Burkina Faso	1	France	47%	4083	France	26%	4083	Côte d'Ivoire	33%	823
Burkina Faso	2	Côte d'Ivoire	19%	823	Côte d'Ivoire	21%	823	France	31%	4083
Burkina Faso	3	United States	6%	7641	**Togo**	5%	762	Belgium	4%	4320
Burundi	1	Belgium	24%	6484	**Kenya**	15%	867	Kenya	19%	867
Burundi	2	France	15%	6372	**Tanzania**	12%	1172	Italy	15%	5316
Burundi	3	Germany	15%	6460	Belgium	11%	6484	Belgium	13%	6484
Cambodia	1	United States	48%	14208	**Thailand**	26%	536	Thailand	32%	536
Cambodia	2	Thailand	10%	536	**Singapore**	17%	1152	China	18%	3351
Cambodia	3	Singapore	10%	1152	Hong Kong, China	16%	1542	Hong Kong, China	15%	1542
Central African Rep.	1	France	48%	5186	France	41%	5186	France	24%	5186
Central African Rep.	2	Spain	9%	4592	**Cameroon**	11%	790	Netherlands	20%	5486
Central African Rep.	3	Cameroon	8%	790	United States	11%	9919	United States	13%	9919
Chad	1	France	50%	4248	France	31%	4248	France	25%	4248
Chad	2	Nigeria	16%	1415	United States	26%	9052	United States	15%	9052
Chad	3	Cameroon	5%	1001	Cameroon	9%	1001	Cameroon	7%	1001
Congo, Dem. Rep.	1	Belgium	21%	6225	**South Africa**	19%	3314	South Africa	19%	3314
Congo, Dem. Rep.	2	Germany	13%	6257	Belgium	18%	6225	Belgium	13%	6225
Congo, Dem. Rep.	3	United States	13%	10270	France	10%	6050	France	9%	6050
Ethiopia(incl. Eritrea)	1	Italy	15%	4473	United States	15%	11221	Saudi Arabia	22%	1930
Ethiopia(incl. Eritrea)	2	Japan	14%	10401	China	13%	8316	China	17%	8316
Ethiopia(incl. Eritrea)	3	Germany	12%	5541	**Saudi Arabia**	11%	1930	United States	9%	11221
Gambia, The	1	United Kingdom	28%	4484	**Senegal**	13%	156	Senegal	17%	156
Gambia, The	2	China	10%	12366	United Kingdom	9%	4484	Côte d'Ivoire	11%	1656
Gambia, The	3	Japan	10%	14027	**Brazil**	8%	5254	Brazil	8%	5254
Ghana	1	United Kingdom	19%	5114	**Nigeria**	10%	423	China	12%	11832
Ghana	2	Germany	14%	5154	United Kingdom	9%	5114	Nigeria	10%	423
Ghana	3	United States	14%	8246	United States	8%	8246	India	6%	8498

Reporter (Importer)		1970 -1974 Partner (Exporter)	Trade share	Distance (km)	2000 -2004 Partner (Exporter)	Trade share	Distance (km)	2005 -2006 Partner (Exporter)	Trade share	Distance (km)
Guinea	1	France	32%	4692	France	16%	4692	China	13%	12652
Guinea	2	United States	17%	6673	Belgium	8%	4953	France	12%	4692
Guinea	3	Italy	10%	4537	United States	8%	6673	United States	8%	6673
Guinea-Bissau	1	Portugal	53%	3059	**Senegal**	26%	364	Senegal	26%	364
Guinea-Bissau	2	United Kingdom	11%	4628	Portugal	23%	3059	Italy	21%	4315
Guinea-Bissau	3	Netherlands	6%	4868	China	8%	12441	Portugal	19%	3059
Haiti	1	United States	59%	2476	United States	59%	2476	United States	55%	2476
Haiti	2	Japan	7%	13116	**Colombia**	3%	1577	Brazil	4%	5452
Haiti	3	Canada	6%	2871	Japan	3%	13116	Colombia	4%	1577
India	1	United States	23%	11762	United States	10%	11762	China	11%	3785
India	2	Japan	12%	5848	Belgium	8%	6420	United States	7%	11762
India	3	United Kingdom	11%	6721	**China**	7%	3785	Saudi Arabia	6%	3053
Kenya	1	United Kingdom	32%	6828	**United Arab Emirates**	13%	3438	India	15%	5439
Kenya	2	Japan	14%	11266	United States	10%	11853	United Arab Emirates	15%	3438
Kenya	3	Germany	12%	6510	**South Africa**	9%	4106	United States	9%	11853
Lao PDR	1	Thailand	45%	525	Thailand	63%	525	Thailand	72%	525
Lao PDR	2	United States	16%	13488	**China**	10%	2779	China	11%	2779
Lao PDR	3	Japan	13%	4144	**Vietnam**	9%	479	Singapore	3%	1863
Liberia	1	Japan	61%	14434	Korea, Rep.	33%	13479	Korea, Rep.	45%	13479
Liberia	2	Norway	11%	6235	Japan	19%	14434	Japan	17%	14434
Liberia	3	Germany	6%	5283	France	15%	4892	Singapore	15%	12721
Madagascar	1	France	53%	8763	France	23%	8763	China	20%	9676
Madagascar	2	Germany	10%	8764	China	14%	9676	France	20%	8763
Madagascar	3	Japan	6%	11423	Hong Kong, China	7%	8573	South Africa	7%	3327
Malawi	1	South Africa	20%	2708	South Africa	45%	2708	South Africa	39%	2708
Malawi	2	United Kingdom	19%	7993	**Mozambique**	8%	1342	Mozambique	14%	1342
Malawi	3	Zimbabwe	19%	522	Zimbabwe	7%	522	Zimbabwe	6%	522
Mali	1	France	38%	4140	France	20%	4140	Senegal	24%	1047
Mali	2	Côte d'Ivoire	15%	927	**Côte d'Ivoire**	18%	927	France	22%	4140
Mali	3	United States	9%	7086	**Senegal**	10%	1047	Côte d'Ivoire	16%	927

		1970 -1974			2000 -2004			2005 -2006		
Reporter (Importer)		Partner (Exporter)	Trade share	Distance (km)	Partner (Exporter)	Trade share	Distance (km)	Partner (Exporter)	Trade share	Distance (km)
Mauritania	1	France	42%	3795	France	21%	3795	France	16%	3795
Mauritania	2	Senegal	14%	422	Belgium	8%	4057	China	8%	11904
Mauritania	3	United States	9%	6036	Spain	7%	2742	United States	8%	6036
Mongolia	1	Czech Republic	48%	6296	China	34%	1172	China	47%	1172
Mongolia	2	Hungary	42%	6193	**Korea, Rep.**	17%	2003	Korea, Rep.	12%	2003
Mongolia	3	Japan	3%	3019	Japan	10%	3019	Japan	11%	3019
Mozambique	1	Portugal	22%	8409	**South Africa**	50%	1622	South Africa	44%	1622
Mozambique	2	Germany	15%	8975	Australia	6%	10739	India	7%	7724
Mozambique	3	United Kingdom	11%	9198	United States	5%	13211	China	5%	11347
Nepal	1	India	59%	801	India	49%	801	India	78%	801
Nepal	2	Japan	15%	5166	**Singapore**	8%	3542	Singapore	3%	3542
Nepal	3	Germany	5%	6859	**China**	6%	3161	Germany	2%	6859
Niger	1	France	41%	3933	France	16%	3933	France	18%	3933
Niger	2	United States	12%	7873	**Côte d'Ivoire**	9%	1133	United States	17%	7873
Niger	3	Germany	9%	4245	United States	9%	7873	Côte d'Ivoire	7%	1133
Nigeria	1	United Kingdom	27%	5025	United Kingdom	11%	5025	United States	12%	8493
Nigeria	2	Germany	14%	5021	United States	10%	8493	United Kingdom	10%	5025
Nigeria	3	United States	12%	8493	France	8%	4722	Netherlands	9%	5110
Pakistan	1	United States	24%	11092	**Saudi Arabia**	12%	2742	Saudi Arabia	11%	2742
Pakistan	2	Japan	11%	5980	**United Arab Emirates**	10%	2084	China	10%	3883
Pakistan	3	Germany	10%	5561	**Kuwait**	8%	2427	United Arab Emirates	10%	2084
Rwanda	1	Belgium	29%	6360	**Kenya**	24%	756	Kenya	29%	756
Rwanda	2	Japan	15%	11914	Belgium	11%	6360	Germany	9%	6331
Rwanda	3	Germany	13%	6331	Germany	9%	6331	Uganda	9%	377
Senegal	1	France	44%	4217	France	25%	4217	France	23%	4217
Senegal	2	Germany	6%	4643	**Nigeria**	14%	2461	United Kingdom	8%	4384
Senegal	3	United States	6%	6157	Thailand	6%	12506	Nigeria	7%	2461
Sierra Leone	1	United Kingdom	28%	4934	United Kingdom	17%	4934	Germany	12%	5129
Sierra Leone	2	Japan	11%	14363	Germany	17%	5129	Côte d'Ivoire	11%	1083
Sierra Leone	3	United States	9%	6947	United States	7%	6947	United Kingdom	8%	4934
Sudan	1	United Kingdom	16%	4942	**Saudi Arabia**	13%	1798	China	23%	8402
Sudan	2	India	13%	4801	China	12%	8402	Saudi Arabia	11%	1798
Sudan	3	China	9%	8402	**United Arab Emirates**	8%	2490	United Arab Emirates	7%	2490

		1970 -1974			**2000 -2004**			**2005 -2006**		
Reporter (Importer)		Partner (Exporter)	Trade share	Distance (km)	Partner (Exporter)	Trade share	Distance (km)	Partner (Exporter)	Trade share	Distance (km)
Tanzania	1	United Kingdom	20%	7505	**South Africa**	14%	3698	South Africa	15%	3698
Tanzania	2	Japan	13%	11404	**Kenya**	9%	677	China	11%	9423
Tanzania	3	Germany	13%	7187	China	8%	9423	India	8%	5667
Togo	1	France	35%	4753	France	19%	4753	United Kingdom	16%	5048
Togo	2	Germany	11%	5069	India	12%	8310	France	16%	4753
Togo	3	United Kingdom	10%	5048	Netherlands	6%	5150	Netherlands	9%	5150
Uganda	1	United Kingdom	26%	6469	**Kenya**	34%	506	Kenya	30%	506
Uganda	2	Japan	13%	11539	**United Arab Emirates**	7%	3570	United Arab Emirates	14%	3570
Uganda	3	Germany	13%	6178	**India**	7%	5691	China	6%	9468
Vietnam	1	United States	46%	13159	**China**	15%	2331	China	20%	2331
Vietnam	2	Singapore	18%	2207	Singapore	14%	2207	Singapore	15%	2207
Vietnam	3	Japan	16%	3673	Japan	14%	3673	Japan	12%	3673
Yemen	1	Japan	19%	9509	**United Arab Emirates**	12%	1466	United Arab Emirates	16%	1466
Yemen	2	Germany	10%	5198	**China**	10%	7417	Saudi Arabia	13%	1063
Yemen	3	Saudi Arabia	10%	1063	United States	8%	11120	India	12%	3691
Zambia	1	United Kingdom	18%	7947	**South Africa**	57%	2279	South Africa	55%	2279
Zambia	2	South Africa	14%	2279	**Zimbabwe**	10%	397	United Arab Emirates	9%	5275
Zambia	3	Japan	9%	12938	**United Arab Emirates**	4%	5275	Zimbabwe	6%	397
Zimbabwe	1	Malawi	36%	522	South Africa	54%	2186	South Africa	17%	2186
Zimbabwe	2	Switzerland	18%	7589	United Kingdom	4%	8293	China	15%	10898
Zimbabwe	3	Germany	14%	8061	Germany	4%	8061	Mozambique	11%	919

Appendix A5. Derivation of the Gravity Model in Panel based on Baier and Bergstrand (2009)

In panel, the estimation of the system of equations (3) and (4) (in main text) raises issues related to estimation of the multilateral trade resistance index and estimation of transport costs. Actually, as we work with panel data, we can no more used fixed effects estimator that includes importer and exporter dummy variables to capture these multilateral resistance indexes (there is no reason to assume that P_i and P_j are constant over time). Then, we follow the simple method for approximating P_i and P_j proposed recently by Baier and Bergstrand (2009) that can be easily extended to a panel framework.[41]

Rewrite the equation (3) and (4) with a temporal dimension:

$$M_{ijt} = \left(\frac{Y_{it}Y_{jt}}{Y_{wt}}\right)\left(\frac{t_{ijt}}{P_{it}P_{jt}}\right)^{1-\sigma} \qquad \text{(A.17)}$$

$$P_{it}^{1-\sigma} = \sum_j \frac{Y_{it}}{Y_{wt}}\left(\frac{t_{ijt}}{P_{jt}}\right)^{1-\sigma} \qquad \text{(A.18)}$$

We apply a first-order log-linear Taylor-series expansion to the system of the 2N multilateral resistance term equations (A.18) (N being the number of countries) to generate a reduced form gravity equation that can be both easily estimated linearly and still consistent with the theoretical model of Anderson and Van Wincoop (2003). The Taylor-serie expansion is "centered" around a World with symmetric trade barriers, i.e. $t_{ijt} = t_t, \forall t$. Under the assumption of bilaterally symmetric trade costs, the solution is:[42]

$$\ln P_{it} = \left[\sum_j \frac{Y_{jt}}{Y_{wt}}\ln(t_{ijt}) - (1/2)\sum_k \sum_j \frac{Y_{kt}}{Y_{wt}}\frac{Y_{jt}}{Y_{wt}}\ln(t_{kjt})\right] \qquad \text{(A.19)}$$

Equation (A.19) represents the GDP-share-weighted (geometric) average of the trade costs facing importer *i* across all exporter *j* in *t*, normalized by the (square root) of GDP-weighted average trade costs.[43] Holding constant (for a given year) bilateral determinant of trade, the largest is *i*'s multilateral resistance, the lower are bilateral trade costs relative to multilateral trade costs and hence the larger the bilateral imports of *i* from *j* in *t* will be.

---

[41] Baier and Bersgtrand (2009) initially proposed this method in a cross-section framework to allow for both efficient estimates of gravity equation and quantitative comparative-static effect without employing the structural system of equations. This approach has been used by Melitz (2008) in his inquiry of the role of language in foreign trade. But we also see in this method an interesting solution for estimating the multilateral resistance term in panel.

[42] This equation corresponds to equation (19) in Baier and Bergstrand (2009, p. 80). See details on the derivation in Baier and Berstrand (2009).

[43] This normalization term is constant, for each year t, across country pairs.

Taking the logarithm of both sides of the augmented trade cost function (14) (in main text):

$$\ln t_{ijt} = \rho_1 \ln(D_{ij}) + \rho_2 \ln(K_{it}) + \rho_3 \ln(K_{jt}) \qquad \text{(A.20)}$$

And then substituting (A.20) in (A.19) yields:

$$\begin{aligned}\ln P_{it} &= \rho_1 \left[\sum_j \frac{Y_{jt}}{Y_{wt}} \ln(D_{ij}) - (1/2) \sum_k \sum_j \frac{Y_{kt}}{Y_{wt}} \frac{Y_{jt}}{Y_{wt}} \ln(D_{kj}) \right] \\ &+ \rho_2 \left[\ln(K_{it}) - (1/2) \sum_k \sum_j \frac{Y_{kt}}{Y_{wt}} \frac{Y_{jt}}{Y_{wt}} \ln(K_{it}) \right] \\ &+ \rho_3 \left[\sum_j \frac{Y_{jt}}{Y_{wt}} \ln(K_{jt}) - (1/2) \sum_k \sum_j \frac{Y_{kt}}{Y_{wt}} \frac{Y_{jt}}{Y_{wt}} \ln(K_{jt}) \right]\end{aligned} \qquad \text{(A.21)}$$

Which can be re-written as:[44]

$$\ln P_{it} = \gamma_t + \rho_1 \left[\sum_j \frac{Y_{jt}}{Y_{wt}} \ln(D_{ij}) \right] + \rho_2 \ln(K_{it}) \qquad \text{(A.22)}$$

with γ_t including all constant terms across country pairs of equation (A.21). Then, substituting $\ln P_{it}$ and $\ln P_{jt}$ in the log-linearized version of (A.17) gives the following equation (corresponding to equation (15) in main text):

$$\begin{aligned}\ln\left(M_{ijt}\right) &= \lambda_t + \beta_1 \ln\left(Y_{it}\right) + \beta_2 \ln\left(Y_{jt}\right) - \theta \ln\left(D_{ij}\right) + \beta_3 \ln\left(K_{it}\right) + \beta_4 \ln\left(K_{jt}\right) \\ &+ \beta_5 \ln\left(MR_{it}\right) + \beta_6 \ln\left(MR_{jt}\right) + \varepsilon_{ijt}\end{aligned} \qquad \text{(A.23)}$$

Where:

$MR_{it} = \left[\sum_k \frac{Y_{kt}}{Y_{wt}} \ln(D_{ik}) \right], MR_{jt} = \left[\sum_k \frac{Y_{kt}}{Y_{wt}} \ln(D_{kj}) \right]$ are multilateral resistance terms,

And:

$\beta_1 = \beta_2 = 1, \theta = \beta_5 = \beta_6 = (\sigma - 1)\rho_1, \beta_3 = 2(\sigma - 1)\rho_2, \beta_4 = 2(\sigma - 1)\rho_3$, λ_t is a vector of year dummies (including then all constant terms across country pairs such as $\ln Y_{wt}$ in (A.17), γ_t in (A.22) or world fuel price index), $\nu_{ijt} = \mu_{ij} + \upsilon_{ijt}$ with μ_{ij} is a specific bilateral random effect, and υ_{ijt} is the idiosyncratic error term with the usual properties.

---

44 If we strictly follow Baier and Bergstrand (2009), one should also introduce, in addition to multilateral distance and infrastructure, other multilateral "border" variables such as adjacency, language, colonial history, etc. Actually, if all these "border" variables, which are constant country-pairs characteristics, are included in the specific bilateral effect, once re-computed as "multilateral" terms they are no longer constant. However, the time variation for all these multilateral "border" variables is only generated by the evolution of the weight i.e. the GDP-share of countries in the World, which is common for all country-pairs.Hence, if introducing these multilateral "border" variables in addition to multilateral "distance" (also invariant except for the the GDP-share) are meaningful in cross-section estimation, they are not relevant in panel, especially with also both bilateral specific effect and year dummies introduced in the equation.