

# The Distance Puzzle and Low-Income Countries: An Update\*

Céline CARRÈRE

Jaime DE MELO

John WILSON

 CÉLINE CARRÈRE, University of Geneva, FERDI and CEPR.

 JAIME DE MELO, University of Geneva, FERDI and CEPR.

 JOHN WILSON, World Bank.

## Abstract

The “distance effect” measuring the elasticity of trade flows to distance has been found 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”. However, this puzzle is regularly challenged by new developments in the specification of the gravity equation or in its estimations. We propose an original survey on the existing methods used to quantify the distance puzzle – basically the computation of an average distance of trade, a meta-analysis on existing gravity papers and the implementation of recent econometric developments, all on a well-specified gravity equation both in cross-section and panel data. We apply all these methods to a unique large database (124 countries from 1970 to 2006). It appears that if all these new developments can change the amplitude of the increase in the trade elasticity to distance, none solve the distance puzzle. We confirm the existence of this puzzle and identify that it only applies to low-income countries who exhibit a significant rising distance effect on their trade of around 18% between 1970 and 2006 while the distance “puzzle” for trade within richer countries disappears.

**JEL Classification:** F10; F40

**Keywords:** International Trade; Gravity Model; Distance Effect

\* This paper is a substantially revised version of Carrère, de Melo and Wilson (2009). We are grateful to Caroline Freund, Nuno Limao, Jacques Melitz, Marcelo Olarreaga, two referees and participants at seminars at the World Bank and CERDI for useful comments on an earlier draft of this paper. This paper has benefited from support from DFID.

## 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”. 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 (e.g. Carrère and Schiff, 2005 or Freund and Berthelon, 2008). Surprising as it may seem, this is consistent with another puzzling result that suggests a burden of distance on bilateral trade that may have, in fact, been increasing over time. 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,  $\theta$ . This seems not to be the case (see e.g. Brun et al. 2005)). This evidence, based on different methods, feeds the debate on the “Distance puzzle”.

In an extensive meta-analysis of 103 gravity model studies, Disdier and Head (2008) estimate an average elasticity,  $\theta$ , of about 0.9. Of course this estimate is likely to capture more than transport costs such as divergence in tastes and preferences, and there are objections to its high value (see Grossman, 1998 and Anderson and van Wincoop, 2004). Exploiting a temporary shock to distance - the closing of the Suez canal in 1967 and its reopening in 1975 - Feyrer (2009) proposes a better-defined estimate of the distance effect on transport costs for trade in goods than those based on the traditional gravity equation. He finds an elasticity of trade with respect to distance of 0.2-0.5, about half as large as the usual estimates.

Although this survey provides several estimates of  $\theta$ , some of which are more plausible than previous ones, this is not our main concern. Rather, we wish to establish the robust result that the *increasing* estimates for  $\theta$  over time are not due to a statistical or econometric misspecification, but to an actual economic phenomenon that remains to be explained. Actually,

the “distance puzzle” is regularly challenged by new developments in the specification of the gravity equation or in its estimation. For instance, Felbermayr and Kohler (2006) show that ignoring zero trade flows when estimating the gravity equation can generate an “artificial” or spurious distance puzzle. In the same vein, Head et al. (2009) claim to solve the distance puzzle thanks to the fixed effects that capture the entry into the sample of distant countries with low trading propensities. The literature that tackles the distance puzzle can therefore be very confusing.

In this paper, we use three approaches to tackle the puzzle. First, we study the evolution of the average distance of trade in the same vein than Carrère and Schiff (2005). Second, using the data in Disdier and Head (2008), we explore the relevance of three explanations of the puzzle (composition effects, treatment of zero trade flows and omitted variable bias) in recent empirical literature. Third we implement recent econometric developments on a well-specified gravity equation both in cross-section and panel data. These methods are applied to this large database (124 countries from 1970 to 2006) that allows us to understand the source of the discrepancies that come out of the recent studies. These different approaches lead us to a clear conclusion on “distance puzzle” debate. We find that, if these new developments can change the amplitude of the increase in the trade elasticity to distance, they do not solve the distance puzzle. Rather, it is composition effects that explain the discrepancy across papers: the distance puzzle is clearly an issue for trade involving the poorest countries. So, depending on the proportion of poor countries in the sample, the distance puzzle is more or less significant.

The three complementary approaches proposed in this paper to assess the distance puzzle focus on the poorest countries as it is precisely this group of countries that should benefit most from a ‘flatter’ world.<sup>1</sup> This focus is also motivated by our earlier work (Brun et al. (2005), Carrère and Schiff (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 analyzes the raw data by constructing an indicator of the evolution of the average distance of trade. 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

---

<sup>1</sup> the third poorest countries in the sample of 124 countries (i.e. the low-income countries according the World Bank classification of 2006).

role of distance in bilateral trade that mainly holds for Low income countries.

Section 3, drawing on the data in meta-analysis by Disdier and Head, reviews the usual explanations for the distance puzzle in the recent gravity literature: composition effects, sample selection (e.g. the problem of the treatment of zero trade flows), the choice of econometric methods, and omitted variable bias (e.g. multilateral resistance terms). We explore the relevance of each of these explanations on the *evolution* over time of the distance effect in a gravity equation. While Disdier and Head find evidence that the recent developments in the gravity literature (both on the theoretical and econometric sides) have significant effects on the *average level* of the distance coefficient, no such effects are found on the *evolution of the coefficient* (the only exception is for papers focusing on developing countries where a very strong distance puzzle is evident after 1970).

Finally, in section 4, we revisit the gravity-predicted  $\theta$  elasticities. To control for as many factors as possible, and to maximize robustness, we carry out both cross-section and panel formulations and use several methods to deal with zero trade flows. This allows us to isolate the effect of different estimators used in the gravity literature on the distance puzzle using still the same extensive sample of 124 countries over 1970-2006. Whatever the estimator used, a distance puzzle is revealed for the same 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 Regionalization of Trade for Low-income countries

### 2.1 Evolution of the average distance of Trade

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:

$$ADOT_t = \sum_i \sum_j \frac{X_{ijt}}{X_{wt}} D_{ij} \quad (1)$$

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 only product of the countries' GDPs (denoted  $Y_{(i)t}$ ):

$$ADOT_t^P = \sum_i \sum_j \frac{X_{ijt}^P}{X_{wt}^P} D_{ij} \quad (2)$$

With  $X^P$  being the potential (or frictionless trade) defined as:

$$X_{wt}^P = \sum_i \sum_j X_{ijt}^P = \sum_i \sum_j \frac{Y_{it} Y_{jt}}{Y_{wt}}$$

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 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 well for developing countries.

Then, we compute the ratio of actual versus potential trade, called the average distance ratio (ADR):

$$ADR_t = ADOT_t / ADOT_t^P \quad (3)$$

The values of these ratios are reported for a sample of 124 countries over the period 1970-2006.<sup>2</sup> To iron out fluctuations, each point is a 5-year average. As suggested by some studies (e.g. Brun et al., 2005 or Carrère and Schiff 2005), we also report averages for the richest and poorest tercile of countries (each tercile has 39 observations). Note that the poorest tercile matches perfectly the low-income country group as defined by the World Bank in 2006.<sup>3</sup> To ease the reading, we set  $ADR_t$  to 1 for the period 1970-1974 (right y-axis in figure 1a, left y-axis in figure 1b).

Figure 1a shows 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. Figure 1b shows a large fall in the average distance of trade for the lowest tercile implying that poor countries have increased trade relatively more with nearby than with distant partners. For this group the average distance of imports (ADOT) fell by more than 15% from 7200 kms in 1970-1974 to 6000 in 2005-2006.

---

<sup>2</sup> 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 countries does not change the results presented here and in the rest of the paper (results available upon request). Nominal bilateral trade flows (in US\$, 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 is taken from the World's Bank World Development Indicators 2008. Distance measure is 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).

<sup>3</sup> 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.

Figure 1. Average distance and Indirect Trade Cost Measures  
124 countries, 1970-2006

Figure 1a – Overall

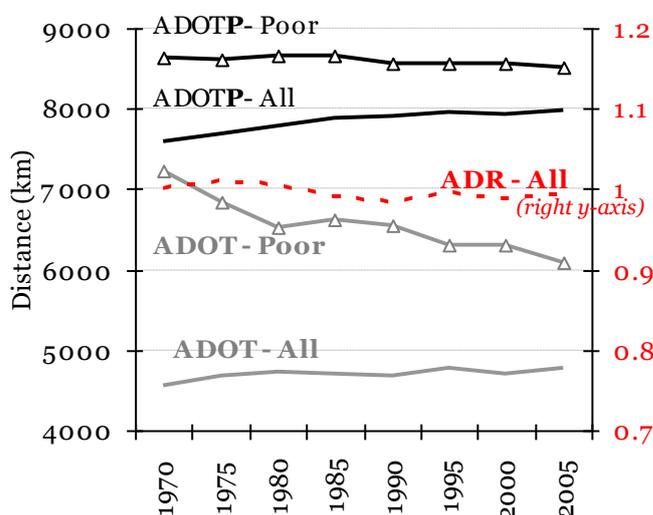
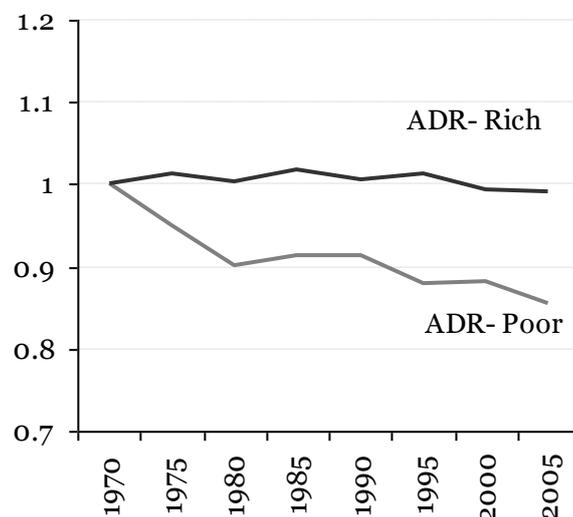


Figure 1b – by Income Tercile



Note: 5-years periods over 1970-2004 and a 2-years period 2005-2006  
Figure 1a: ADOT values should be read on the left-y-axis while ADR refers to the right-y-axis.  
Source: authors' calculations on data from UN-COMTRADE and WITS.

Then, there is clear evidence of diverging paths of the ADRs for the whole sample and for the lowest tercile: 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 ratio over the sample period.

## 2.2 Extensive versus intensive margins

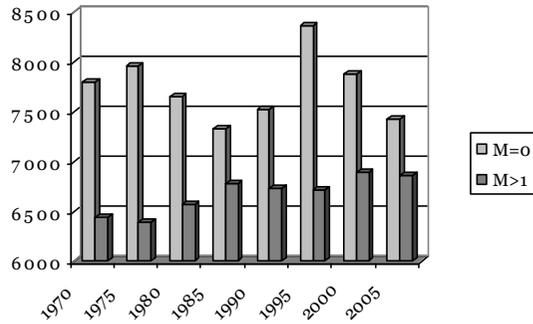
Why did poor countries trade relatively more with geographically closer partners? <sup>4</sup> In this simple setting, the only two possibilities are changing weights of existing trading partners, or changing trading partners. First, it could be that close trading partners (e.g. China and India in Asia) 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 1a indicates that the potential distance of trade barely increases, so this effect cannot be a major factor.

<sup>4</sup> 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 2 Average distance of zero trade flows, Richest and Poorest Terciles.

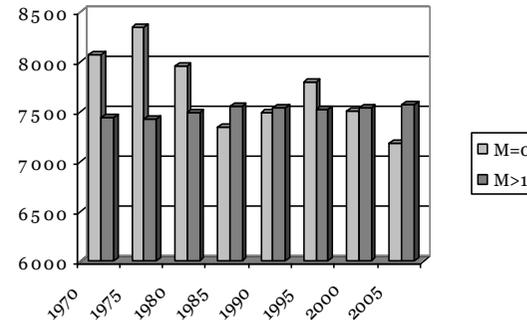
Poorest Tercile (39)

Average Distance (km, unweighted)



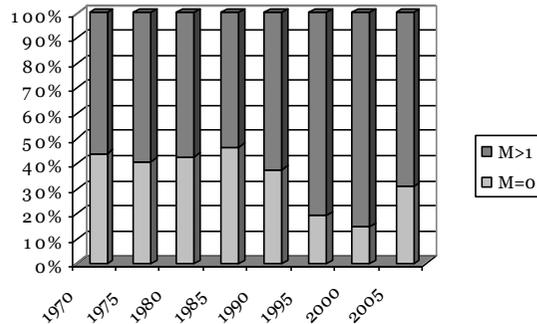
Richest Tercile (39)

Average Distance (km, unweighted)



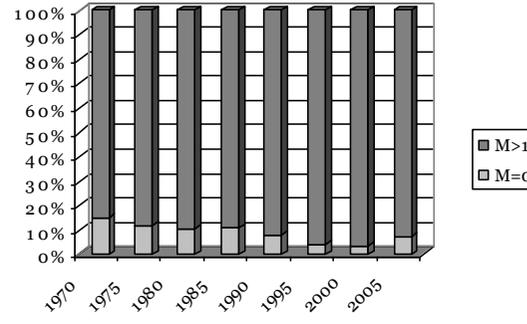
Poorest Tercile (39)

Number of observations (balanced=4,797)



Richest Tercile (39)

Number of observations (balanced=4,797)



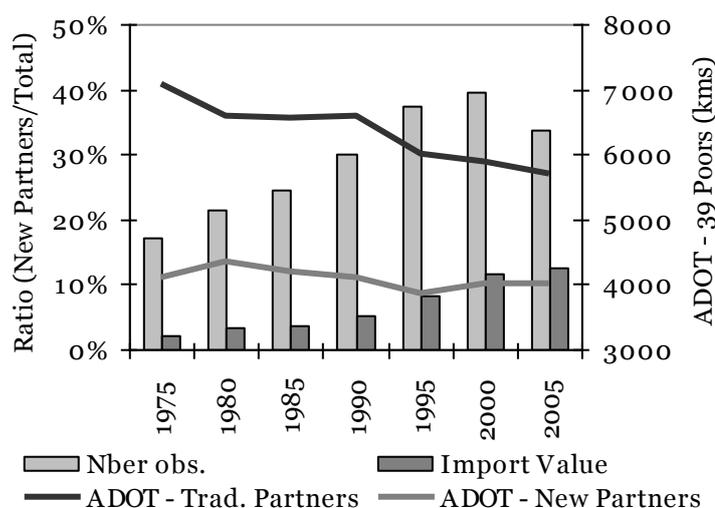
Note: 5-years periods over 1970-2004 and a 2-years period 2005-2006

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 2, 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.<sup>5</sup> Then the number of zero trade flows decreases sharply falling 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 3 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 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 3: 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

<sup>5</sup> 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 or generate new trade with geographically closer partners.<sup>6</sup> It is clear from figure 3 that most of the regionalization of trade is generated by trade redistribution within the intensive margin.

The conclusion from this inspection of the raw data is that the poorest countries have increased their trade share with geographically closer partners which would be expected 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. This would be consistent with recent literature on “sequential exports” (e.g. Albornoz et al., 2010): the new exporters attempt to learn about their export profitability by first entering close foreign markets (corresponding to lower entry costs). 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.

Next sections explore if this increasing elasticity of trade to distance only applies for the poorer countries in the sample after controlling, in a gravity framework, for some of the factors that could alter distance-sensitive trade costs. We first propose a meta-analysis based on existing estimates (section 3) and then re-estimate the elasticity of trade to distance in our unified sample of 124 countries from 1970 to 2006 (section 4).

### **3. THE DISTANCE PUZZLE IN THE GRAVITY LITERATURE**

#### **3.1. The Rising Distance Effect in the gravity equation**

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,

---

<sup>6</sup> 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 3 are unchanged (available upon request).

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%].<sup>7</sup> This range is typical of cross-section (sometimes averaged over 5-year periods) estimates of aggregate trade volumes where trade costs,  $t_{ij}$ , are given by:<sup>8</sup>

$$t_{ij} = (D_{ij})^\rho \prod_{m=1}^M (z_{ij}^m)^{\gamma_m} \quad (4)$$

where the set  $z_{ij}^m$  ( $m=1, \dots, M$ ) 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:

$$\ln(M_{ij}) = \alpha_0 + \alpha_1 \ln(Y_i) + \alpha_2 \ln(Y_j) + \theta \ln(D_{ij}) + \sum_{m=1}^M \lambda_m \ln(z_{ij}^m) + \varepsilon_{ij} \quad (5)$$

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 Anderson and van Wincoop (AvW, 2004), and as illustrated in the above typical estimates of  $\tilde{\theta}$ , these appear to be implausibly high.<sup>9</sup> This leads 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  $z_{ij}^m$ .

---

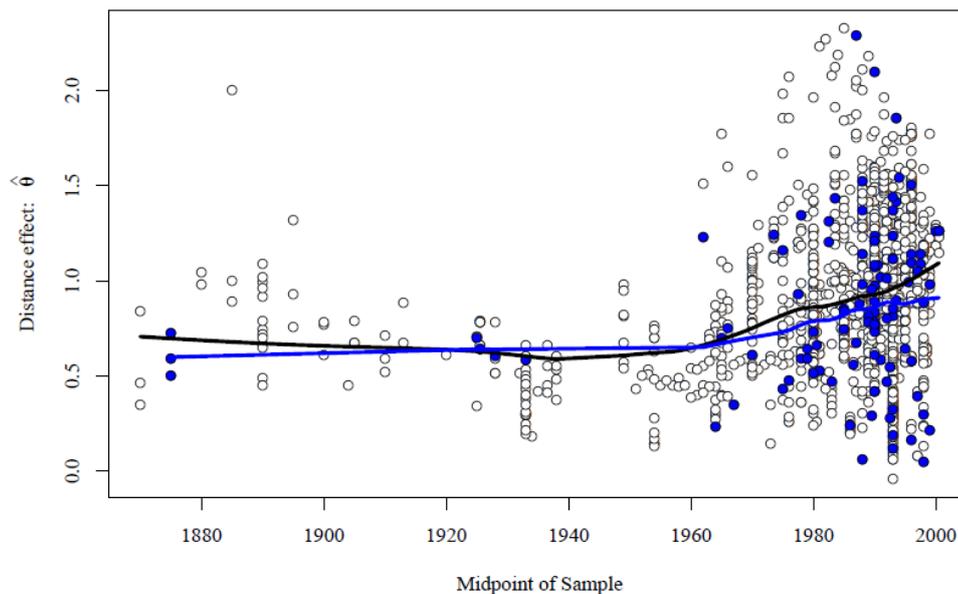
<sup>7</sup> The general formula is:  $M_1 / M_0 = (D_1 / D_0)^{\tilde{\theta}}$

<sup>8</sup> We use a function to proxy trade costs as direct overall trade costs measure since “direct measures are remarkably sparse and inaccurate” (Anderson and van Wincoop, 2004, p.692). Direct measures are only available for a few components, for instance transportation and insurance costs, usually proxied by the ratio of c.i.f. and f.o.b. trade values (with all induced problems, see Hummels and Lugovskyy, 2006). An interesting micro-founded measure of bilateral trade frictions is proposed by Chen and Novy (2010) but it is based on the ratio of domestic versus bilateral trade flows. To be used, it requires production data on gross flows which are not available for poor countries.

<sup>9</sup> In the theory-based gravity equation, the elasticity of trade to trade costs depends on the elasticity of substitution in consumption,  $\sigma$  according to  $\alpha = (\sigma - 1)$ . Since  $\sigma$  has to be estimated separately (as reported by AvW,  $5 < \sigma < 10$ ), 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.

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 4 below reproduced from the 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).<sup>10</sup> This figure plots the elasticity estimates against time and fits a kernel smoother through the data (dark line). Based on  $\sigma=8$ , they estimate that the overall border barriers to trade amount to around 50%. From their survey of estimates reported in figure 4 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 4. The rising Distance Effect in Gravity Models



<sup>10</sup> They built their base sample from English language papers listed in Econlit (78 papers, 60 of which are published in academic journals, 4 are chapters in books, and 14 are working papers) and JSTOR databases (25 additional papers). Then they deleted estimates that were not in the form of elasticities and some extreme outliers using the Grubbs test. After the above deletions, the 103 studies provided 1467 usable observations. Their online appendix, available at [http://strategy.sauder.ubc.ca/head/papers/meta\\_papers.pdf](http://strategy.sauder.ubc.ca/head/papers/meta_papers.pdf), lists the full sample of 103 papers, including the number and range of estimates from each paper. They span a relatively large period going from 1870 to 2001, including 188 pre-1970 sample estimates.

Note: 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. y-axis: Distance effect ( $\hat{\theta}$ ) / x-axis: Midpoint of Sample

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

To give an idea of the orders of magnitude suggested by the meta-analysis in figure 4, distance impedes trade by 37% more since 1990 than 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 (but regularly challenged), is referred to as the “distance puzzle” or the “missing globalization puzzle” (Coe et al. 2007).

### 3.2. The Gravity model Set-up

Even though most estimates in figure 1 come from the “traditional” gravity equation in (5), 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 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 (an extensive review is proposed by Anderson, 2010).<sup>11</sup> Here is one. Take a one sector economy with a representative consumer with CES preferences with common elasticity  $\sigma$  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  $p_{ij} = p_i t_{ij} = p_i (1 + \tau_{ij})$  and trade costs enter

<sup>11</sup> See also 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  $\sigma$  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.

multiplicatively as in (5). 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:

$$M_{ij} = \left( \frac{Y_i Y_j}{Y_w} \right) \left( \frac{t_{ij}}{P_i P_j} \right)^{1-\sigma} \quad (6)$$

$$P_i^{1-\sigma} = \sum_j \frac{Y_j}{Y_w} \left( \frac{t_{ij}}{P_j} \right)^{1-\sigma} \quad (7)$$

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  $j$ ,  $\sigma > 1$  is the elasticity of substitution in the CES utility function. According to (6) and (7), 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 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.<sup>12</sup> 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  $\tau_{ij}$ . The reduced-form distance-dependent trade cost function would read:

$$t_{ij} = (1 + \tau_{ij}) (D_{ij})^\rho \prod_{m=1}^M (z_{ij}^m)^{\gamma_m} \quad (8)$$

where the ad-valorem equivalent of trade costs includes all border trade costs, depend on product characteristics and on the market characteristics

---

<sup>12</sup> 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%.

of the transport sector.<sup>13</sup> In practice, the functional form of trade costs  $t_{ij}$  is in fact given by (4). Substituting (4) into (6) and (7), the estimated equation in a cross-section setting becomes:

$$\ln(M_{ij}) = -\ln(Y_w) + \ln(Y_i) + \ln(Y_j) - (\sigma - 1)\rho \ln(D_{ij}) + \sum_{m=1}^M (\sigma - 1)\gamma_m \ln(z_{ij}^m) + (\sigma - 1)\ln(\tilde{P}_i) + (\sigma - 1)\ln(\tilde{P}_j) + \varepsilon_{ij} \quad (9)$$

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. Estimates of (9) can be computed from several methods, but most of them are obtained by the inclusion of country fixed-effects which in addition to 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) include these country fixed-effects. As before, a “distance puzzle” obtains when the distance effect,  $\theta$ , takes larger values when estimated from more recent data as exemplified in figure 4.

Economic and econometric arguments have been advanced to explain the presence of the puzzle in a gravity equation. 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 evolution estimates of  $\theta$  to these three set of controls, using the meta-analysis dataset of Disdier and Head (2008).

### 3.3. Explanations for the sensitivity of Distance Elasticity estimates

*Composition Effects.* Composition effects appearing through the elasticity of substitution at the product level, have been invoked most.<sup>14</sup> Two recent

---

<sup>13</sup> 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  $t_{ij} = [(f_{ij} + \tau_{ij})(D_{ij})]^\rho + \sum_{m=1}^M \gamma_m z_{ij}^m$  where  $f_{ij}$  is the freight rate.

<sup>14</sup> Disdier and Head (2008) discuss three channels that will yield different estimates of  $\theta$  in theory-based gravity models: differences in  $\sigma$  across products, differences in the

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 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.<sup>15</sup>

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

---

response of trade costs to distance across goods  $\rho$  and differences in productivity across firms (see also the discussion in AvW, 2004, p.726).

<sup>15</sup> 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  $\theta$  neither when we introduce country-pair specific effects that control for difference in latitude for instance.

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.<sup>16</sup> 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 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, 2009), 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

---

<sup>16</sup> 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.

flows) are closer to the “true” values. In other words, if zero trade flows are positively correlated with distance (which is clearly the case as discussed above), then ignoring zero trade flows when estimating the gravity equation can generate an “artificial” or spurious distance puzzle. However, as shown by Helpman, Melitz and Rubinstein (2008) and Larch et al. (2010), in addition to the selection bias arising from omitting zero trade flows, another bias emerges if the heterogeneity of firms is not controlled for, as estimation captures the distance effect on both intensive and extensive margins. And this bias can be large, larger than the omitted zero trade flows. This joins the findings of Bernard et al. (2007) based on US export for 2000, who show that aggregate trade relationships are heavily influenced by extensive-margin adjustments both in terms of the number of destinations and the number of exported products. However, as we note in section 4, disentangling between extensive and intensive margins in a gravity equation requires having good identification variables (i.e. variables that allow to identify the probability to exports versus the amount traded) over quite a long time period which is a difficult exercise.

*Multilateral terms and others.* Relatedly, omitting the multilateral terms when estimating (9) 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  $\varepsilon_{ij}$  (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.

### **3.4. Sensitivity of Distance Elasticity estimates in a meta-analysis**

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:

$$\tilde{\theta}_{ij} = \alpha_0 + \alpha_1 D^{70-79} + \alpha_2 D^{80-89} + \alpha_3 D^{90-99} + \sum_m \beta_m x_{ij}^m + u_i + e_{ij} \quad (10)$$

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). Finally, the  $u_i$  are random effects. In (10), positive estimates for  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  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, remains. In sum, the meta-analysis persists in showing a rising estimate of  $\tilde{\theta}$  across samples, specifications and econometric methods.

However, thanks to the estimations of equation (10), Disdier and Head can conclude that the increasing distance effect after 1970 observed in their meta-analysis is not due to, for instance, the increasing number of developing countries in the samples of more recent papers. But what about the evolution of the distance coefficient *within* papers that all include developing countries? To see this, we extend their exploration by interacting the controls,  $x_{ij}^m$  with time dummies, i.e. we estimate:<sup>17</sup>

$$\begin{aligned} \tilde{\theta}_{ij} = & \alpha_0 + \alpha_1 D^{70-79} + \alpha_2 D^{80-89} + \alpha_3 D^{90-99} \\ & + \alpha_1^m \left( D^{70-79} x_{ij}^m \right) + \alpha_2^m \left( D^{80-89} x_{ij}^m \right) + \alpha_3^m \left( D^{90-99} x_{ij}^m \right) + \sum_m \beta_m x_{ij}^m + u_i + e_{ij} \end{aligned} \quad (11)$$

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

---

<sup>17</sup> We thank Anne-Célia Disdier and Keith Head for sharing their database.

theory i.e. by including either remoteness variables or country fixed effects.

Coefficients of interest are reported in figures 5a-5c. 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 coefficient 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: the estimates 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).<sup>18</sup> 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).

Figure 5. Distance Puzzle in existing empirical literature

Figure 5a

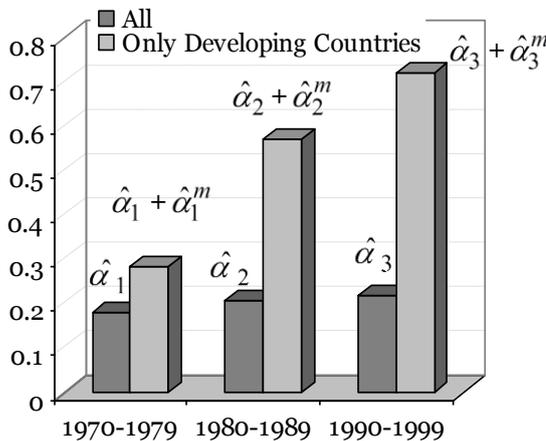
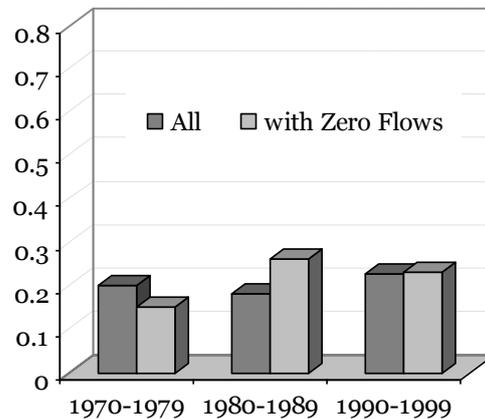


Figure 5b



<sup>18</sup> For instance, over the 1467 point estimates, only 52 (from 7 different studies) concerns developing countries.

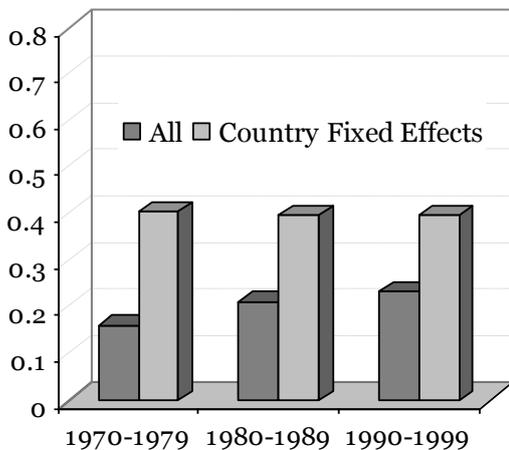


Figure 5c.

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

#### 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 poorest countries in the sample. This is indeed what comes out of the alternative estimates below: repeated cross-sections (each cross-section representing a 5 years average) with 2 dimensions – importer and exporter countries- as in the vast majority of cases reported in the meta-analysis; and panel estimates, i.e. with 3 dimensions (importers, exporters and time) 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 (8) 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 and then turn to the panel estimates.

##### 4.1 Cross-section Estimates

We follow the by-now standard approach and estimate:

$$\ln(M_{ij}) = \alpha_0 + \alpha_i + \alpha_j + \theta \ln(D_{ij}) + \sum_m \lambda_m \ln(z_{ij}^m) + v_{ij} \quad (12)$$

where  $\alpha_i$  and  $\alpha_j$  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.<sup>19</sup>

$D_{ij}$  is the distance between  $i$  and  $j$  and  $z_{ij}^m$  includes dummy variables indicating whether the two countries are contiguous, share a common language, or have had a common colonizer.<sup>20</sup> In this first stage, we do not include the usual preferential trade agreement (PTA) dummy as this is not a structural but a policy-based variable that can be both an explanatory factor and a result of the “distance puzzle” (see section 4.3).

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.<sup>21</sup> 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  $\ln(1 + M_{ij})$  instead of  $\ln(M_{ij})$  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 ( $\ln(1 + M_{ij})$ ). However, the OLS estimator does not take into account the censorship of the dependent variable. Column (3) reports the results from the Eaton and Tamura (ET, 1994) tobit with  $\ln(a + M_{ij})$  as dependent variable. Under the ET estimator, instead of arbitrarily imposing  $a = 1$ , the value of the  $a$  parameter is endogenously determined and the dependent

---

<sup>19</sup> 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) and discussed in Anderson (2010) will be used in the panel estimates reported below. We will also propose a panel regression that introduces time-varying exporter and importer fixed effects (it and jt).

<sup>20</sup> From the Centre d'Etudes et de Prospectives et d'Informations Internationales (CEPII).

<sup>21</sup> All results are reported in appendix A3, table A3.1.

variable will be censored at the value  $\ln(a)$ .<sup>22</sup> 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.<sup>23</sup> 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.<sup>24</sup> 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.<sup>25</sup>

---

<sup>22</sup> 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).

<sup>23</sup> We fit conditional fixed effects PPML (see details in Arellano and Honoré, 2001, section 5).

<sup>24</sup> See discussion in Martin and Pham (2008) and Santos-Silva and Tenreyro (2009).

<sup>25</sup> 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 (HMR) (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. Larch et al. (2010) propose to use the HMR estimator to solve the distance puzzle but they use “common language” and “common religion” variables as instruments in the first stage. However, these instruments present two main caveats: (i) there are time-invariant (contrary to the export costs entry that we want to proxy and that can actually influence the distance puzzle through their evolution) and (ii) they are dubious as these variables are often highly significant in a gravity equation and then can be suspected to violate the exclusion restriction. Another possibility is to achieve identification by functional form assumptions (non-linearity of the selection equation), using exactly the same set of regressors in the selection and the outcome equations (see e.g. Deb and Trivedi, 2006, for an application of this method in panel). However, in our case, the identification strategy is not sufficient as the outcome equation (the gravity estimation on positive trade flows controlling for the “mills” ratio) gives very similar results to those in OLS on  $\ln(M)$ , with an increase in the (absolute value of the) distance coefficient of around +30% at the intensive margin during the period. Note that in the first stage probit (extensive margin), the impact of distance on the probability to trade is decreasing (in absolute value) over time.

Table 1. Barriers to Trade: Cross-section results (124 countries, 1970-2006)

Methods dependent var.	OLS	OLS	ET- Tobit	PPML	OLS	OLS	ET- Tobit	PPML
	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)
	<b>1970</b>				<b>2005</b>			
lnDij	-1.066*** (0.0352)	-1.115*** (0.034)	-1.231*** (0.034)	-0.668*** (0.000)	-1.429*** (0.0303)	-1.336*** (0.038)	-1.325*** (0.029)	-0.710*** (0.000)
Common Border	0.898*** (0.142)	0.326* (0.178)	0.244* (0.132)	0.367*** (0.000)	0.821*** (0.146)	0.613*** (0.188)	0.584*** (0.115)	0.555*** (0.000)
Common Language	0.702*** (0.0638)	1.001*** (0.058)	1.201*** (0.063)	0.328*** (0.000)	0.967*** (0.0585)	1.112*** (0.064)	0.967*** (0.053)	0.203*** (0.000)
Colonial links	1.315*** (0.114)	1.401*** (0.134)	1.329*** (0.161)	0.589*** (0.000)	0.639*** (0.109)	0.663*** (0.142)	0.695*** (0.141)	0.00188*** (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 <sup>2</sup> or pseudo-R <sup>2</sup>	0.698	0.745	0.743	0.877	0.799	0.817	0.910	0.905
t-student on distance coeff. 2005 versus 1970					7.82***	4.33***	2.14**	542.17***

Standard errors in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Fixed country effects are not reported.

Source: authors' calculations

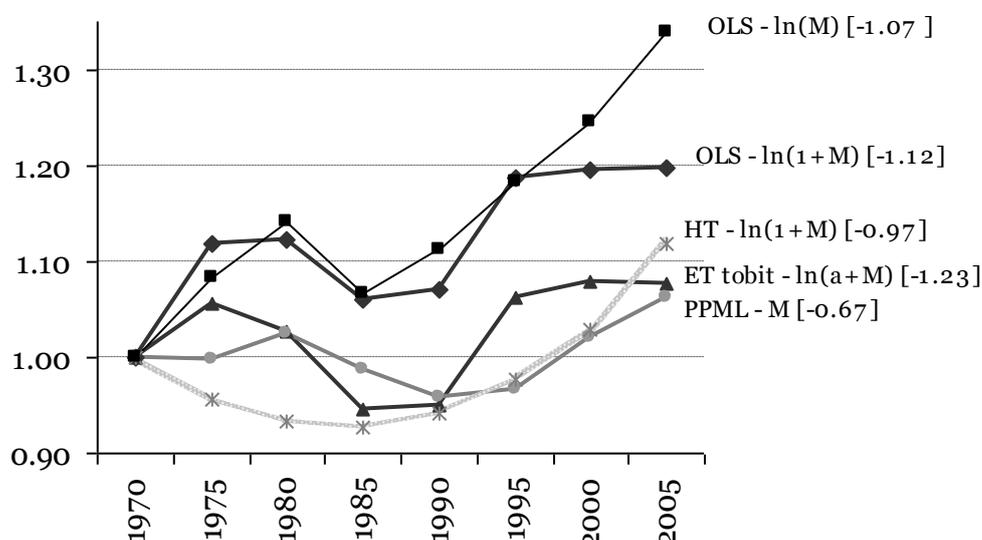
Several patterns stand out in table 1. First, as expected, the dummies have the usual signs and usual significance levels. There are some changes in the coefficient values over the sample period reflecting expectations from a globalizing world. The value of the common language coefficient falls drastically 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. However, we also observe a clear and significant increasing impact of distance on trade, coupled with an increasing importance of sharing a common border in the 3 last specifications. Both confirm the regionalization of trade emphasized earlier.

The estimates of  $\theta$  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.<sup>26</sup> The explanatory

<sup>26</sup> 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.

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.<sup>27</sup> One reason for this better fit would be 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  $\theta$



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  $\theta$  in figure 6 (with all distance coefficients normalized to unity on the first sub-period 1970-1974) confirms the existence of a puzzle.<sup>28</sup> 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

<sup>27</sup>  $R^2$  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 pseudo- $R^2$  are based on ratios of log-likelihoods.

<sup>28</sup> The distance coefficient for 2005 is always significantly lower than the one for 1970, at 1% level for the 2 OLS regressions and for the PPML and at 5% for the ET tobit (see the last line in table 1).

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  $\theta$  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:

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

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.

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  $\theta$  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  $\theta$  increases in the range [+18%; +19.5%] 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 45% in the PPML regressions).

Figure 7. Cross-Section Estimates of  $\theta$  by Tercile

Figure 7a. 39 poorest' trade with all countries

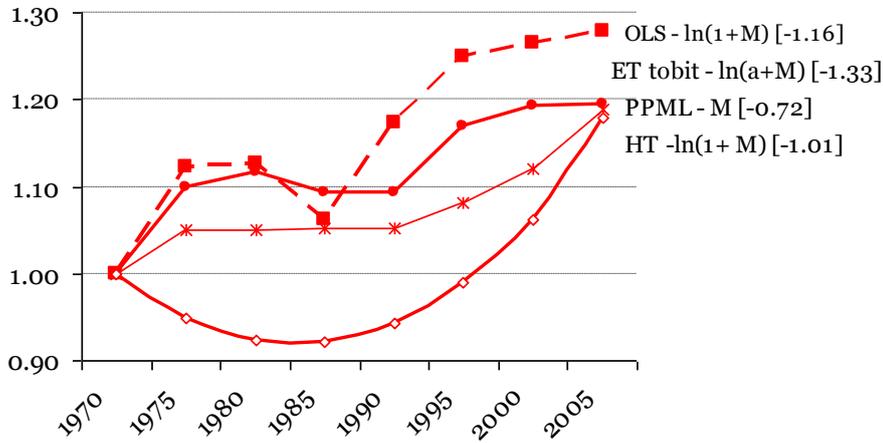
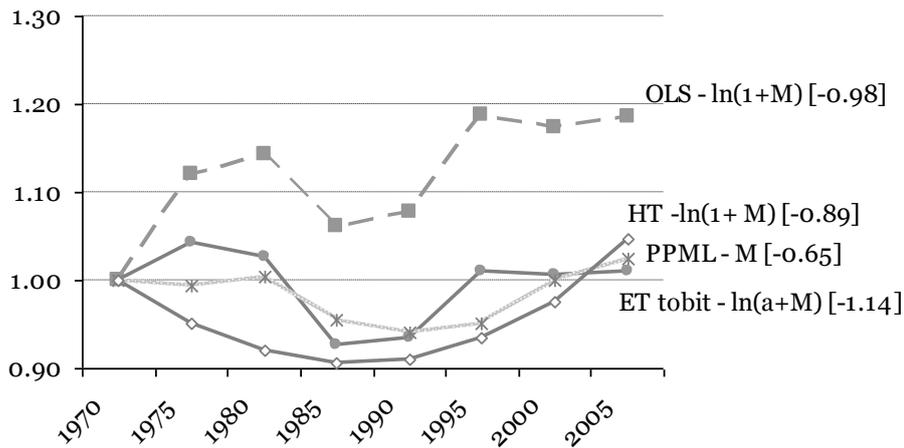


Figure 7b. Within 39 richest countries



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

## 4.2. Panel Estimates

Panel estimates allow the introduction of trade costs that vary through time and that are lumped together in (8) under the term  $\tau_{ij}$  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:<sup>29</sup>

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

As constructed (see appendix A2), larger values for  $K_{i(j)t}$  indicate a better infrastructure.<sup>30</sup> 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.<sup>31</sup>

This trade cost function is introduced in (6) and (7) 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  $\theta$  by half over the period 1970-1995.<sup>32</sup> Furthermore, the

---

<sup>29</sup> 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. We maintain  $D_{ij}$ , our variable of interest, as the use of random effects (see below) allow us to estimate the corresponding coefficient (but not the fixed effects specification). Nor does trade costs function include fuel costs due to the year dummies.

<sup>30</sup> 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).

<sup>31</sup> See Brun et al. (2005, appendix D).

<sup>32</sup> 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% (see OLS with country fixed effects on  $\ln(M)$  in figure 6). Once he controlled for North-South distance,

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).<sup>33</sup> 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:

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

where

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

or

“remoteness” variables,<sup>34</sup> and  $\lambda_t$  is a vector of year dummies,  $\lambda_{ij}$  the bilateral fixed effects (with  $\lambda_{ij} \neq \lambda_{ji}$ ) and  $\varepsilon_{ijt}$  the error term.<sup>35</sup> A quadratic time trend:

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

is introduced in(15) to detect any significant evolution of  $\hat{\theta}$ . Because estimation in panel with bilateral fixed effects (FE) drops  $\theta_1$ , 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 (

Melitz obtained an increase in distance of around +8%, while with our corresponding random effect HT specification on  $\ln(M)$  over 1970-1995 (with implicitly control for North-South distance as for all time-invariant country-pair characteristics) yields an estimated increase of +3.1% (computation based on col 4 table A3.2).

<sup>33</sup> 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.

<sup>34</sup> 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.

<sup>35</sup> Note that the random effect specification implies  $\varepsilon_{ijt} = \mu_{ij} + v_{ijt}$  with  $\mu_{ij}$  is a specific bilateral random effect, and  $v_{ijt}$  is the idiosyncratic error term with the usual properties.

$K_{i(j)t}$ ) are correlated with the bilateral random effects ( $\mu_{ij}$ ), we correct for this endogeneity using the Hausman-Taylor (1981) instrumental variables method.<sup>36</sup>

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 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 (2) 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 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.

As a robustness check, we also regress equation (15) using time-varying importer and exporter fixed effects as done for instance by Head et al. (2009). Results reported in column (3) confirm the significance and the magnitude of the distance puzzle.<sup>37</sup>

---

<sup>36</sup> 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.

<sup>37</sup> Head et al. (2009) claim to solve the distance puzzle thanks to the introduction of these importer\*time and exporter\*time fixed effects that capture “the entry into the sample of distant countries with low trading propensities”. However, when looking at their results reported in table 4 page 437, the distance puzzle disappears when they confine the sample to observations where *both* services and trade flows are non-missing and non-zero, i.e. to

Table 2. Barriers to Trade: Panel estimations (124 countries, 1970-2006)

Methods		FE	HT	FE
dependent var.		ln(1+Mijt)	ln(1+Mijt)	ln(1+Mijt)
		(1)	(2)	(3)
GDP it	<i>lnYit</i>	0.754*** (0.031)	0.763*** (0.009)	-
GDP jt	<i>lnYjt</i>	0.906*** (0.032)	0.914*** (0.009)	-
Distance ij	<i>lnDij</i>	-	-0.989*** (0.024)	-
Multil. Resistance it	<i>lnMRit</i>	2.075*** (0.253)	1.609*** (0.051)	-
Multil. Resistance jt	<i>lnMRjt</i>	1.573*** (0.234)	1.209*** (0.051)	-
Infra it	<i>lnKit</i>	0.208*** (0.029)	0.191*** (0.010)	-
Infra jt	<i>lnKjt</i>	0.176*** (0.033)	0.159*** (0.010)	-
trend in Dist	<i>tlnDij</i>	0.00961*** (0.002)	0.0111*** (0.001)	0.0102*** (0.000)
trend <sup>2</sup> in Dist	<i>t<sup>2</sup>lnDij</i>	-0.000376*** (0.000)	-0.000377*** (0.000)	-0.000313*** (0.000)
<i>Bilateral specific effects</i>		<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>Year dummies</i>		<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>(reporter*year) &amp; (partner*year) fixed effects</i>		<i>no</i>	<i>no</i>	<i>yes</i>
N. obs		417,110	417,110	417,110
N. country pairs		12,432	12,432	12,432

Robust standard errors in parentheses after corrections for clustering of country pairs in col (1) and (3) and using the correction proposed in the Hausman and Taylor model in col. (2); \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

a sample of 64 high or upper middle income countries. Hence, their results are in line with our main result, namely that the distance puzzle remains only for poor countries.

### 4.3. Further robustness tests

We carried out four further robustness tests, two on the trend estimates in table 2 (not reported here) and two on the interpretation of the evolution of the distance coefficient for Low-Income countries in both panel and cross-section estimates.

First, we checked the sensitivity of estimates in table 2 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).<sup>38</sup> We do not use the same proxy as Melitz (2007), i.e. the distance North-South, since we already have country-pair specific effects in the equation. 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 effect, these results confirm those in table 2, column 3 where the three sets of fixed effect, (*ij*), (*it*) and (*jt*), control for a large range of explanations than can be found in empirical literature. In sum, the robustness checks from the panel results which allow for bilateral (and asymmetrical) effects do not alter the result that the elasticity of trade to distance has increased for the lower-income countries.

We next check whether the Low-Income distance effect here is not a Sub-Saharan Africa (SSA) specificity since three quarters of the Low-Income (LI) countries are in SSA (see table A1.1). To check this, we introduce an interactive dummy for SSA countries with distance in equation (13). The cross-section PPML estimation for sub-periods 1970-1975 and 2005-2006 (with country fixed effects) reveals that the distance effect for SSA LI countries increases by 27% while for other LI countries it increases by 14%. This heterogeneity in the LI country group is confirmed by the Hausman-Taylor estimation in panel: +22% for SSA-LI countries and +11% for other LI countries. There seems to be an African specificity even if the distance effect increase for other LI countries is far from negligible. However, it is difficult to disentangle the LI from the SSA effect since, as noted above, 30 countries out of 39 in the LI group are from SSA (and 30 out of the 36 SSA countries in the sample are also LI countries).

---

<sup>38</sup> Berthelon 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)).

Finally, we check for a potential effect coming from partnership in a Regional Integration Agreement (RIA), though this effect is difficult to discern in the data since most LI countries are involved in multiple RIAs. We add a PTA dummy in equation (13) and run the cross-section PPML estimation for sub-periods 1970-1975 and 2005-2006 (with country fixed effects), the strong distance effect for the LI group vanishes.<sup>39</sup> However, this result is difficult to interpret since, during this long period, virtually all LI countries reported in table A1.1 participate in at least one PTA with their neighbors. Then, the PTA dummy is also likely to capture other factors contributing to the regionalization of world trade. For instance, participation by LI countries in the Trade Facilitation initiative could contribute to a relative decrease of fixed cost *vs.* variable costs (i.e. distance dependent costs) as administrative costs related to trade decrease. This effect captured by the PTA dummy could, in turn explain the observed regionalization of trade. Moreover, the implementation of a PTA can be endogenous to the regionalization of trade! In the end, the PTA dummy only confirms an increasing trade with nearby 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 focused 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  $\theta$  increasing in the range [+18%; +19.5%] 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 of this regionalization of trade for the low-income (LI) countries could be explored. As shown by Berthelon and Freund (2008), products in some industries are becoming more substitutable making

---

<sup>39</sup> We use the PTA dummy variable available on the J. Bergstrand homepage, see <http://www.nd.edu/~jbergstr/DataEIAs2006/EIA5yData2006.htm>

distance loom larger as small changes in costs will lead to larger shifts in the sourcing of trade.<sup>40</sup> Regionalization of trade might also reflect the implementation of the multiple PTAs that have proliferated among LI countries. We also suggested that this regionalization could reflect the dramatic decrease in a host of costs independent of distance ( MFN tariffs, border-related costs, administrative costs, communication costs or increasing containerization), all of which would enhance the relative importance of transport costs that depend on distance.

The multiple possibilities underlying the observed regionalization of LI trade does not allow one to draw welfare implications. For example, regional trade integration resulting from a preferential tariff reduction for neighboring countries can generate either beneficial trade creation or harmful trade diversion reflected in both new trade flows with nearby countries and/or a switch of trade from distant towards geographically closer partners.

The regionalization of trade could also reflect ‘deep’ integration effects as administrative and technical barriers to trade are being reduced more rapidly for the LI country group relative to others over the period, generating new trade flows that are welfare-increasing. For example, a reduction in trade frictions in LI countries could provide an incentive to move from the informal sector to the formal sector or from the previous formal sector in home trade to the one engaged in foreign trade. This would promote foreign trade generally, but because of the persistence of transport costs in foreign trade, would especially favor foreign trade with close trading partners. If so, this welfare-increasing regionalization of world trade would be captured by the gravity model. Then the indirect evidence (since we do not have time-series data on the evolution of trade costs) in this paper would be good news as it would mean a deepening integration of this group of countries into the World Trading System.

A less optimistic view can be drawn if one assumes that, over the period, a growing part of world trade is generated by vertical specialization and just-in-time production. In this case, trade costs can be viewed as a growing impediment in the supply-chain production. Then, if low-income countries’ trade costs (in particular distance-dependent, such as high markups in international shipping) remain high compared to other developing countries’ trade costs, the observed regionalization of trade could be interpreted as a marginalization of these countries.

---

<sup>40</sup> in a gravity model derived from fundamentals, the distance coefficient is one minus the elasticity of substitution.

## References

- Aidt, T. and M. Gassebner (2008), "Do Autocratic States Trade less?" Mimeo, University of Zurich.
- Albornoz F., Calvo Pardo H.F., Corcos G. and E. Ornelas (2011), "Sequential Exporting", CEPR Discussion Paper No. DP8103
- Anderson, J. (2010). "The Gravity Model". Essay prepared for *Annual Review of Economics*, vol. 3.
- 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.
- Allen, D. and B. Shepherd (2007) "Barriers to Entry, Trade Costs, and Export Diversification in developing Countries", Mimeo, World Bank.
- 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.
- Bernard Andrew B., J. Bradford Jensen, Stephen Redding and Peter K. Schott (2007), "Firms in international Trade", CEP discussion Paper n°795.
- 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 C. and M. Schiff (2005), "On the Geography of Trade: Distance is Alive and Well", *Revue Economique*, 56, 1249-74.
- Carrère C., J. de Melo and J. Wilson (2009) "The Distance Effect and the Regionalization of Trade of Low-Income Countries" CEPR# 7458
- Chen, N. and Novy, D. (2010) "Gravity, Trade Integration, and Heterogeneity across Industries", *Journal of International Economics*, forthcoming.
- 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: Commercial Policy, Monetary Policy, and Corporate Governance, MA: Edward Elgar Publishing Inc., 63-76.
- Disdier A.C. and Head, K. (2008), "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *The Review of Economics and Statistics*, Vol. 90, No. 1. pp.37-41.
- Deb P. and P.K. Trivedi (2006), "Specification and Simulated Likelihood Estimation of a Non-Normal Treatment-Outcome Model With Selection: Application to Health Care Utilization. *Econometrics Journal*, Vol. 9, No. 2, pp. 307-331, July.
- 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, A., Brian F., and A.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.
- Feyrer J. (2009), "Distance, Trade and Income – The 1967 to 1975 closing of the Suez canal as a natural experiment", NBER working paper 15557, forthcoming in *Quarterly Journal of Economics*.
- 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.
- Head, K., Mayer T. and J. Ries (2009), “How remote is the offshoring threat?” *European Economic Review*, 53, 429-444.
- 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. and V. Lugovskyy (2006), “Are Matched Partner Trade Statistics a Usable Measure of Transportation Costs?” *Review of International Economics*, 14(1), 69–86
- Hummels, D., V. Lugovskyy, A. Skiba (2009), “The Trade Reducing Effects of Market Power in International Shipping” *Journal of Development Economics*, forthcoming.
- Larch, M., Norback, P.-J. and D. Urban (2010), “Globalization and the Distance Puzzle”, mimeo CISifo
- Leamer, E. and J. 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.
- Santos Silva, J.M.C and S. Tenreyro . (2009). "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator". Department of Economics, University of Essex, Discussion Paper No 66.
- 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.



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



**Contact**

[www.ferdi.fr](http://www.ferdi.fr)

[contact@ferdi.fr](mailto:contact@ferdi.fr)

+33 (0)4 73 17 75 30