DOCUMENT DE TRAVAIL DT/2012-09

Heterogeneity and the Distance Puzzle

Elizaveta ARCHANSKAIA
Guillaume DAUDIN

UMR DIAL 225
Place du Maréchal de Lattre de Tassigny 75775 • Paris Cedex 16 •Tél. (33) 01 44 05 45 42 • Fax (33) 01 44 05 45 45
• 4, rue d'Enghien • 75010 Paris • Tél. (33) 01 53 24 14 50 • Fax (33) 01 53 24 14 51
E-mail : dial@dial.prd.fr • Site : www.dial.prd.fr

HETEROGENEITY AND THE DISTANCE PUZZLE[1]

Elizaveta Archanskaia
KU Leuven
Sciences-Po-OFCE
archanskaia@gmail.com

Guillaume Daudin[2]
PSL, Université Paris-Dauphine,
LEDa, UMR DIAL, 75016 Paris, France
Sciences-Po, OFCE, 75007 Paris, France
guillaume.daudin@dauphine.fr

Document de travail UMR DIAL

Avril 2014

Abstract:

This paper shows that declining exporter-specific product heterogeneity can explain the non-decreasing distance elasticity of trade in 1963-2009. The paper first examines common explanations of the distance puzzle: sample and sectorial composition effects and the rise of FTAs. In the Armington framework, perceived increasing substitutability of exporter specific product bundles, i.e. the elasticity of trade flows to trade costs, can explain an increase in the distance coefficient. We provide robust empirical evidence that was the case over 1963-2009. Consequently, the well-documented increase in the distance coefficient is compatible with a reduction in the elasticity of trade costs to distance.

Key words: Gravity equation, distance puzzle, trade elasticity, trade costs.

Résumé

Cet article montre que la diminution de l'hétérogénéité des variétés de produits par chaque exportateur peut expliquer l'absence de réduction de l'élasticité du commerce à la distance entre 1963 et 2009. L'article examine tout d'abord des explications courantes de ce « paradoxe de la distance » : l'effet d'échantillon, l'effet de composition et la montée des traités de libre-échange. Dans le cadre d'Armington, l'augmentation de la substituabilité subjective entre les paniers de biens produits par chaque exportateur, autrement dit l'élasticité des flux commerciaux aux coûts du commerce, peut expliquer l'augmentation du coefficient de la distance. Nous fournissons des données empiriquement robustes pour montrer que cela a été le cas entre 1963 et 2009. En conséquence, l'augmentation du coefficient de la distance est bien compatible avec une baisse de l'élasticité des coûts du commerce à la distance.

Mots Clés : Equation de gravité, paradoxe de la distance, élasticité du commerce, coûts du commerce.

JEL Code: F15, N70

[1] This paper has greatly benefited from helpful discussions with Thomas Chaney, Anne-Célia Disdier, Peter Egger, Lionel Fontagné, Raphaël Franck, Joseph Francois, Guillaume Gaulier, Samuel Kortum, Jacques Le Cacheux, Philippe Martin, Thierry Mayer, and suggestions by seminar participants at the ETSG, the AFSE, the Ljubljana Empirical Trade Conference (Forum for Research in Empirical Internation Trade), the international workshop on the Economics of Global Interactions (Bari), at the Economics Department of SciencesPo, and at the International Trade Working Group at the University of Chicago. We thank Christian Broda, Natalie Chen, Dennis Novy, and David Weinstein for giving us access to their programs. The usual disclaimers apply.

[2] Corresponding author (guillaume.daudin@dauphine.fr).

Introduction

The estimated effect of distance in gravity equations has unambiguously increased over the last 60 years. Disdier and Head (2008) adopt a meta-analytical approach and find that distance impedes trade by 37% more in the 1990s than it did from 1870 to 1969[1]. Head and Mayer (2013) estimate the distance elasticity of trade in successive cross-sections and find that it has doubled in 1960-2005. This increase in the distance elasticity of trade has been dubbed the "distance puzzle", as the common opinion is that technological developments in transportation and communication, e.g. the airplane, the container, and the internet, would have led to the "death of distance" by the end of the 20$^{\text{th}}$ century[2].

This paper investigates the empirical relevance of the possibility that within the Armington framework the non-decreasing distance elasticity of trade is due to an increasing sensitivity of consumers to price differences (i.e. a reduction in the perceived heterogeneity of country-specific goods bundles). We make the simple point that the flattening out of the world may go hand in hand with a persistent impeding effect of distance on trade if consumers perceive product bundles shipped out by each country to the world market as increasingly substitutable.

Recent work has seeked to rationalize the distance puzzle in three complementary ways: by pointing out a possible misspecification of the econometric model, by refining the specification of the trade cost function, and, more recently, through the lens of network analysis.

The first strand of the literature has investigated the incidence of the estimation method on the magnitude of the distance puzzle. Santos Silva and Tenreyro (2006) advocate estimating the gravity model in multiplicative form using a specific non-linear estimator, the Poisson Pseudo Maximum Likelihood (PPML). Contrary to the canonical log-linear approach, this estimator provides consistent coefficient estimates and is robust to rounding error and overdispersion which are both likely features of trade data[3]. The magnitude of the distance puzzle is reduced when the gravity model is estimated in multiplicative form. Thus, Bosquet and Boulhol (2009) find that the distance elasticity stays within the .6-.75 range between 1948 and 2006. It is hence more accurate to state the puzzle as a non-decreasing distance elasticity of trade since the 1960s.

The sensitivity of the distance puzzle to the estimation method is likely due to sample com-

[1] See also Berthelon and Freund (2008); Combes et al. (2006); Brun et al. (2005); Buch et al. (2004).
[2] Cairncross (1997); Levinson (2006); Friedman (2007)
[3] Santos Silva and Tenreyro (2011) and Fally (2012) provide evidence on desirable properties of the PPML. Inadequacy of alternative non-linear estimators is discussed in Bosquet and Boulhol (2009, 2010). Head and Mayer (2014) review properties of alternative estimators.

position effects. Head and Mayer (2013) show that the magnitude of the puzzle in the log-linear specification is reduced in the sample of stable trade partners. Indeed, the growth of trade has been both intensive in the sense that the volume of established trade relations has increased and extensive in the sense that new trade relations have been established (Helpman et al. (2008); Baldwin and Harrigan (2011)). If trade relations have in priority been established between small and distant partners, the reduction in the number of zeros may have gradually reduced the underestimation of the distance coefficient in the loglinear specification[4]. This explanation echoes Felbermayr and Kohler (2006) who pointed out that the log-linear specification was subject to sample selection bias due to the exclusion of zero trade flows. They conjectured that the distance puzzle was an artefact of reduction in this bias through the extensive margin of trade[5].

The second and most prominent strand of the literature singled out the underpinnings of the trade cost function as key to understanding the distance puzzle. The basic point formulated by Buch et al. (2004) is that the distance elasticity of trade is invariant to reductions in transportation and communication costs if their distribution over distance remains unchanged Furthermore, while the distance elasticity of transport costs may have decreased (Hummels (2007)), other cost components, such as delays, may have become more distance-elastic (Hummels and Schaur (2013)). More generally, if freight costs account for an ever smaller fraction of distance-dependent trade costs, the distance elasticity of trade will be determined by other, possibly persistent, cost components[6].

A complementary mechanism is proposed in Krautheim (2012) in the heterogeneous firms' framework. He models the informational component of trade costs as a fixed cost which decreases in the number of exporting firms. This refinement of the trade cost function magnifies the distance elasticity of trade because the number of exporters is decreasing in variable trade costs which increase with distance. This magnification mechanism may have been reinforced by the increasing weight of information costs in total fixed costs, explaining increasing elasticities.

An alternative explanation put forward in models with input-output linkages is that the relationship between total trade costs and transport costs may be non-monotonic. An increasing

[4] Using a non-linear estimator which controls for the mass of exporting firms Larch et al. (2013) find a decreasing distance elasticity in 1980-2006. They attribute the puzzle to the growing bias of the OLS estimate.

[5] This leaves open the question of the estimator which correctly captures the level of the distance elasticity. Head and Mayer (2013) argue that PPML gives too little weight to small trade flows characteristic of more distant partners. For Santos Silva and Tenreyro (2006) small trade flows are more prone to measurement error.

[6] Head and Mayer (2013) propose a typology of persistent but unobserved trade costs. Anderson and Wincoop (2004) provide a decomposition of total trade costs. Daudin (2003, 2005) put forward that trade costs may have remained stable as a share of value added.

distance elasticity may be an endogenous outcome of transport cost reductions if they engender a reoptimization of the production process which ends up increasing the relative cost of long-distance trade. One possible mechanism is trade cost magnification through multiple border crossings by goods as a consequence of increased production fragmentation (Yi (2010); Daudin et al. (2011); Noguera (2012)). Another mechanism formalized by Duranton and Storper (2008) works through quality upgrading. Lower transport costs shift trade towards higher-quality inputs which are more distance-sensitive because their customization requires intensive communication, e.g. more back-and-forth travelling, between upstream and downstream firms.

The focus of the literature on the shape of the trade cost function mirrors the expectation that the distance coefficient moves together with the elasticity of trade costs to distance. But Chaney (2013) provides a theoretical foundation for the gravity equation through the lens of network analysis which demonstrates that the distance coefficient can be invariant to the trade cost function. In this model the rate of distance decay in aggregate trade is linked to the rate of decay in the density of firms which cover that distance with their network of contacts. As the geographic dispersion of the network is increasing in firm size, the shape parameter of the firm size distribution plays a key role in explaining movements in the distance coefficient. Thus, technological advances in transportation increase the geographic dispersion of exports at the level of the firm but have no incidence on the distance elasticity of aggregate trade as long as the stationary firm size distribution verifies Zipf's law.

The link between the distance coefficient and the parameter which captures the degree of structural heterogeneity in the economy is not specific to Chaney (2013). Every theoretical foundation of the gravity model delivers a functional relationship of the distance elasticity with the intensity of the incentive to trade, e.g. the degree of structural heterogeneity in some model-specific dimension. The combination of empirical evidence on the changing shape of the trade cost function with evidence on the stability of the distance distribution of trade indicates that structural heterogeneity may have contributed to the evolution of the distance coefficient. However, empirical evidence on the evolution of structural heterogeneity in the economy since the 1960s is notoriously scarse (Head and Mayer (2013)).

We pursue the idea that a key parameter for understanding movements in the distance coefficient is the one measuring the degree of structural heterogeneity in the economy. Following Arkolakis et al. (2012), we refer to this parameter as the 'trade elasticity'. We offer an estimation method in the Armington framework in which structural heterogeneity captures the degree of product differentiation by place of production (i.e. product heterogeneity). We estimate the distance elasticity of trade and the Armington trade elasticity in each year between 1963 and 2009 and deduce the implied evolution of the elasticity of trade costs to distance[7]. Our main result is that the increase in the Armington trade elasticity not only rationalizes the non-decreasing distance elasticity of trade but also hints at a reduction in the elasticity of trade costs to distance between 1963 and 2009.

To the best of our knowledge, Berthelon and Freund (2008) is the only paper which has investigated the impact of changes in perceived product substitutability on the distance coefficient. Using estimates of sectoral Armington elasticities obtained by Broda and Weinstein (2006), Berthelon and Freund (2008) find a positive relationship between the variation in sectoral distance coefficients and the variation in Armington elasticities between 1985-1989 and 2001-2005[8]. Our approach is different from Berthelon and Freund (2008) because we focus on the aggregate Armington elasticity and provide direct estimates of this parameter in each year between 1963 and 2009. Our approach is complementary to Broda and Weinstein (2006) because instead of investigating the degree of differentiation among sectoral varieties we document the increasing similarity of the product mix that countries supply to world markets.

The paper proceeds in three steps. First, we examine the hypothesis that the distance puzzle is a by-product of compositional changes in the set of trading pairs or in the set of traded goods, or in the rise of FTAs. Second, we suggest a straightforward demand estimation procedure Third, we present the results of this method and conduct a number of robustness checks. We discuss the biais introduced by the presence of zero trade flows. We address endogeneity concerns by instrumenting unit values (our price proxy) with the real exchange rate which is specific to each bilateral relationship. We find robust empirical evidence that the trade elastcity has increased faster than the distance elasticity of aggregate trade between 1963 and 2009. The evolution of the distance coefficient is thus compatible with an decrease of the elasticity of trade costs to distance.

[7] Erkel-Rousse and Mirza (2002) do this exercise for just one point in time on a subsample of world trade flows.

[8] Berthelon and Freund (2008) work with 776 sectors defined at the SITC Rev.2 4-digit level. Broda and Weinstein (2006) use time-series variation in prices and market shares for the set of exporters to the US market to get one value for the Armington elasticity in 1972-1988 and another value in 1990-2001.

1 The magnitude of the distance puzzle

In this section we evaluate the sensitivity of the distance puzzle in 1963-2009 to composition effects identified as explanatory of movements in the distance coefficient in previous estimations of the loglinearized gravity model. In particular, Head and Mayer (2013) find that the distance puzzle is reduced in the loglinear specification in the balanced sample between 1960 and 2005 while Berthelon and Freund (2008) find that changes in the sectoral composition of world trade are not explanatory of movements in the distance coefficient in 1985-2005. We find that the distance puzzle is magnified in the sample of stable pairs and robust to fixing the product composition of world trade when the model is estimated in multiplicative form. We also check whether the conduct of trade policy helps rationalize the distance puzzle. If Free Trade Agreements (FTAs) reduce the relative cost of within-FTA trade and FTA formation takes place at short-distance, regional integration would result in an increasing intensity of within-FTA trade, and mechanically induce an increasing distance elasticity of trade costs. This however, is subject to endogeneity concerns.

1.1 The magnitude of the sample composition effect

We use the COMTRADE dataset to make our investigation of the distance puzzle directly comparable to Head and Mayer (2013) and Berthelon and Freund (2008). We work with the 4-digit SITC Rev.1 product classification (600-700 goods) because it provides the longest and most comprehensive coverage of disaggregate bilateral trade (1963-2009)[9]. Data on bilateral distance, bilateral trade cost controls such as adjacency, common language, colonial linkages, and data on belonging or having once belonged to the same country are taken from the CEPII[10].

We restrict the sample to trade in goods which are attributed to specific 4-digit categories and to pairs for which we have data on bilateral trade cost controls. App. A lists the resulting set of countries. For each active pair attributed sectoral flows are summed to obtain total bilateral trade. The resulting sample covers between 88% and 99% of reported trade in COMTRADE [11].

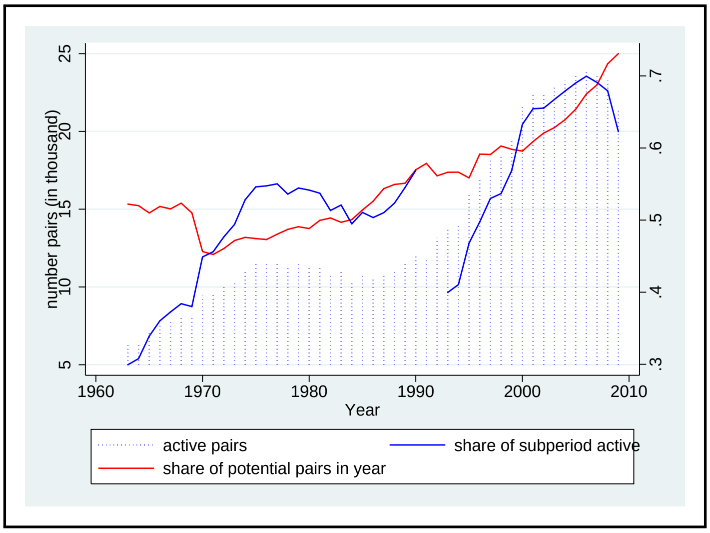
[9] The earliest available year is 1962. The estimation is conducted on import flows. We focus on 1963-2009 for consistency with the timeframe used in estimation of Armington trade elasticities in section 3.

[10] See Mayer and Zignago (2011). The database is available at www.cepii.fr. We constructed bilateral distance and bilateral cost controls for East and West Germany, USSR, and Czechoslovakia.

[11] COMTRADE itself covers between 70% to 90% of total world merchandise trade according to the WTO (http://stat.wto.org/Home/WSDBHome.aspx, accessed in May 2011). As trade data is of better quality for imports, the estimation is conducted on import flows.

Fig.1 summarizes the coverage of the data. The number of active pairs increases more than fourfold in 1963-2009 (in dash, left scale), both because more countries report trade to COMTRADE and because more pairs have non-zero trade flows (Helpman et al. (2008)). Active pairs make up between 45% and 70% of the total number of possible trade relationships, with a clear upward trend (in red, right scale).

Figure 1: Active pairs in COMTRADE (1963-2009)



If we focus on the set of pairs reporting non-zero trade in at least one year of the sample, the share of active pairs increases by 20 percentage points between 1963 and 1990 and by 20 percentage points between 1993 and 2009 (in blue, right scale)[12]. By the end of the sample about 2/3 of pairs which trade at least once between 1963 and 2009 are reporting non-zero trade. Hence, sample composition effects are substantial. Nonetheless, the bulk of total trade is attributable to the 1056 pairs which trade both ways in every year. We refer to this set of stable reciprocal pairs as the 'superbalanced' sample and use it to investigate the magnitude of the sample composition effect[13].

We follow the canonical Anderson and van Wincoop (2003) derivation of the gravity model to express aggregate bilateral trade X_{ij} as a function of bilateral trade barriers τ_{ij}, multilateral trade resistance terms in source i and destination j (resp. Π_i and P_j), and nominal incomes Y_n with $n \in \{i, j, w\}$ where w is world income[14].

[12] We split the sample in two subperiods, 1963-1990 and 1993-2009, to take into account country creation and disappearance in the early 1990s.

[13] The superbalanced sample includes 37 countries listed in App. A. The appendix discusses trade coverage.

[14] This formulation is not specific to the Armington framework. See footnote 20 in Eaton and Kortum (2002) and subsequent discussions of equivalence in Arkolakis et al. (2012); Head and Mayer (2013).

$$X_{ijt} = \left(\frac{Y_{it}Y_{jt}}{Y_w t}\right)\left(\frac{\tau_{ijt}}{\Pi_{it}P_{jt}}\right)^{\varepsilon_t} \tag{1}$$

We include the time subscript t not only on each variable but also on the elasticity of trade flows to trade costs ε to underline that this parameter is subject to change. In the Armington framework $\varepsilon_t = 1 - \sigma_t$ where σ_t corresponds to the elasticity of substitution between goods of different national origin. We seek to quantify the evolution of the elasticity of aggregate trade flows to trade costs: σ_t is the key parameter of interest for this paper.

As total bilateral trade costs τ_{ijt} are not directly observed for each pair and year, we model them as a function of observable time-invariant bilateral controls which are distance, adjacency, and common language together with persistent but time-varying controls standard in the gravity literature which are historical and current colonial linkages as well as belonging or having once belonged to the same country. We include an unobserved bilateral trade cost component ν_{ijt} assumed to have mean zero conditional on the observables[15]. We denote distance $\eth_{ij}$, group the other time-invariant observables in the vector Z and time-varying observables in the vector S_t to get the following specification of the trade cost function:

$$\tau_{ijt} = \exp\left\{\rho_t \ln \eth_{ij} + Z'\zeta_t + S_t{}'\varsigma_t + \nu_{ijt}\right\} \tag{2}$$

Replacing (2) in (1), substituting source and destination specific variables with country fixed effects (resp. f_{it} and f_{jt}), defining a constant ξ_t and specifying a multiplicative error term ξ_{ijt} which includes the exponentiated unobserved bilateral trade cost gives the equation to be estimated on aggregate bilateral trade:

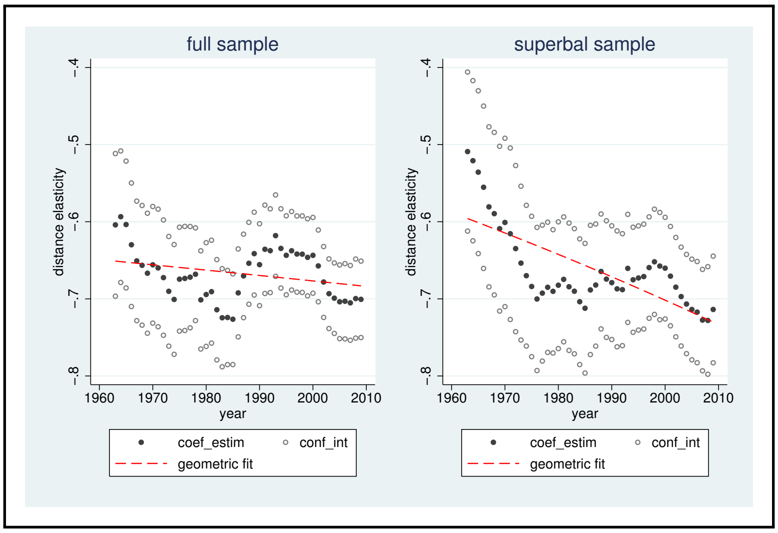
$$X_{ijt} = \exp\left(\xi_t - \delta_t \ln \eth_{ij} + Z'\zeta_t + S_t{}'\varsigma_t + f_{it} + f_{jt}\right)\xi_{ijt} \tag{3}$$

To ensure consistency of the point estimates we do not loglinearize the model although switching to a non-linear estimator may entail a loss of efficiency (Manning and Mullahy (1999)). We implement (3) in the full and superbalanced samples using the PPML estimator (Santos Silva and Tenreyro (2006)). The estimation is conducted in cross section. The parameter of interest is the distance elasticity, $-\delta_t$, which corresponds to the product of the distance elasticity of trade costs ρ_t and of the trade elasticity ε_t[16].

[15] This assumption can be questioned, in particular with respect to excluded trade cost controls which vary as a result of trade policy decisions. We examine this question in 1.3.

[16] We follow notation in Head and Mayer (2013).

Figure 2: The sample composition effect



Results for both samples are shown in fig.2. In terms of the point estimates the distance sensitivity of trade has increased by 4.9% in the full sample between 1963-2009 (left pane). This increase is magnified to 22.5% in the superbalanced sample (right pane)[17]. There is indeed a distance puzzle, and it is exacerbated in the sample of stable trade relationships[18].

Computing heteroskedasticity-robust standard errors yields the large confidence intervals shown in fig.2, putting in doubt even the actual existence of the distance puzzle. This is not the case when usual standard errors are used. Considering that the PPML approach already takes into account heterskedasticity, and that we are looking at these point estimates as if they were descriptive statistics on the whole population, we do not believe that invalidates our approach.

1.2 The magnitude of sectoral composition effects

The incidence of sectoral composition effects is assessed in two ways. The first exercise consists in fixing the sectoral composition of world trade. The second exercise consists in fixing the sectoral composition of the bundle supplied by each exporter to the world market.

In the first exercise we fix the sectoral composition of total trade to the initial year of the sample. Denoting each 4-digit sector k and the annual share of the sector in world trade (w,t) by $s^k_{w,t}$, the reweighting procedure fixes the share of each 4-digit sector in world trade to its share

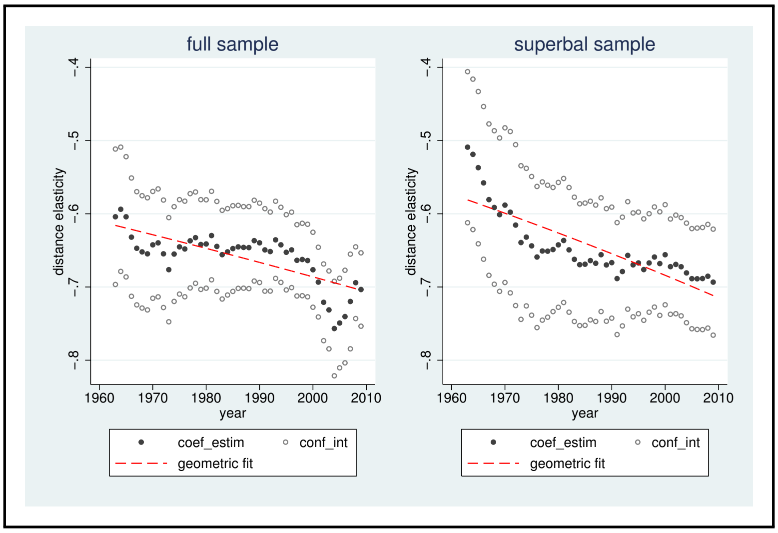
[17] The underlying annual growth rate is 0.10% in the full, and 0.44% in the superbalanced sample.

[18] The opposite result holds in the loglinear specification (see Head and Mayer (2013)). In OLS the decrease in the distance elasticity due to the elimination of small trade flows between distant partners trumps the increase in the distance elasticity of stable trade relationships.

in 1963. The reweighted sectoral bilateral flow is $\tilde{X}^k_{ijt} = X^k_{ijt} * \frac{s^k_{w,1963}}{s^k_{w,t}}$. The reweighted sectoral flows are summed for each pair, and the gravity equation is estimated in each year for aggregate bilateral trade.

Results are shown in fig.3. The evolution of the distance coefficient is much linear with time in the full sample (left pane), and this exacerbates the distance puzzle to a 14% increase in δ_t. The impact of this reweighting procedure is about nil in the sample of stable pairs (right pane) where δ_t increases by 22%[19]. Indeed, the main incidence of fixing the sectoral composition of world trade is the elimination of short-term fluctuations in the distance coefficient due to fluctuations in the weight of the energy sector. As this sector plays a relatively minor role in trade of stable reciprocal partners, the reweighting procedure has little incidence on the distribution of trade over distance in this sample.

Figure 3: Product composition effect: fixing the world bundle

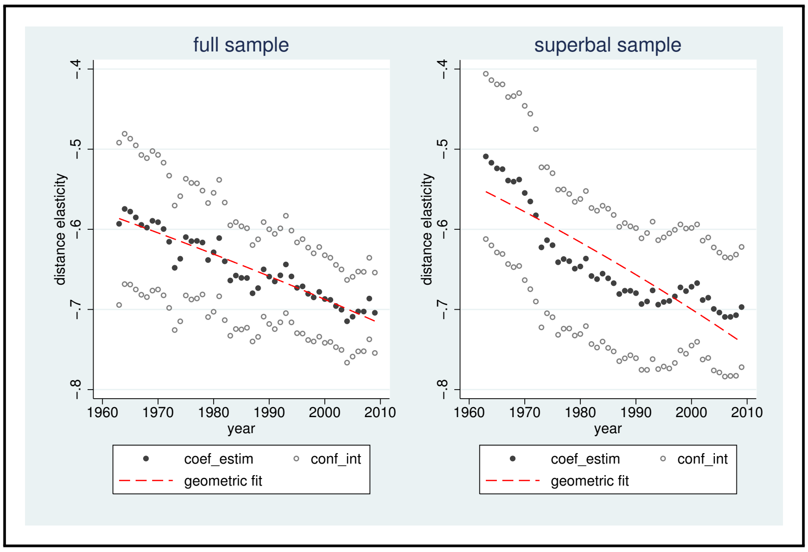


In the second exercise we fix the composition of the bundle supplied by each exporter i to the world market. Denoting the annual share of the sector in world imports from i by $s^k_{i,t}$, the reweighting fixes the share of each 4-digit sector in world imports from i to its share in 1963. The reweighted sectoral bilateral flow is $\tilde{X}^k_{ijt} = X^k_{ijt} * \frac{s^k_{i,1963}}{s^k_{i,t}}$. The resulting sectoral flows are summed for each pair, and the gravity equation is estimated on total reweighted bilateral trade.

As illustrated in fig.4, fixing the composition of the country-specific composite good exacerbates the magnitude of the distance puzzle. Furthermore, while the degree of precision in the estimation of the gravity equation is similar to the benchmark specification, 89% of the varia-

[19] The underlying annual growth rate is 0.29% in the full and 0.44% in the superbalanced sample.

Figure 4: Product composition effect: fixing the country bundle



tion in the distance coefficient is attributable to the time trend in the full sample, against just 8% in the benchmark specification[20]. The corresponding annual growth rate is .43% in the full, and .64% in the stable sample. This corresponds to a 22% increase in the distance sensitivity of trade between 1963 and 2009 in the full sample, and to a 34% increase in the stable sample.

Short-term fluctuations in the distance coefficient are likely to be attributable to product and sample composition effects. But the long-term evolution of the distance elasticity appears linked to structural changes which are reinforced in the set of stable trade relationships.

1.3 The magnitude of the FTA effect

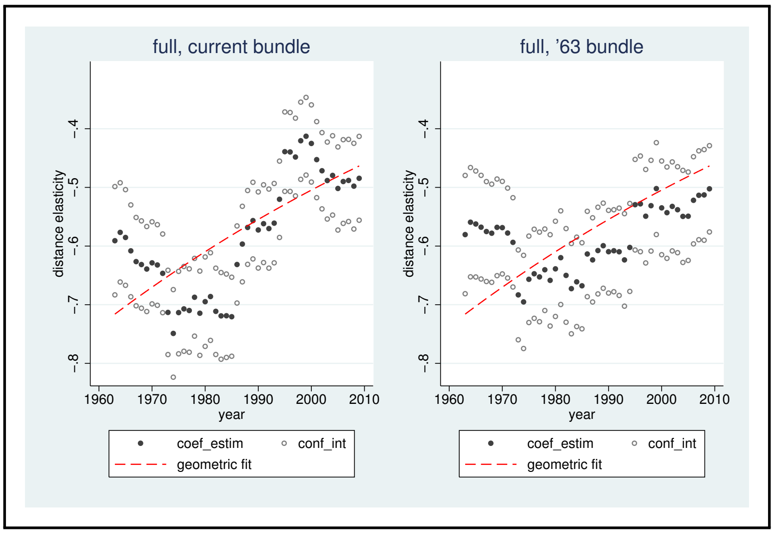
It is possible that policy changes affected the distance distribution of trade costs inducing an increase in the distance elasticity. The distance elasticity of ad valorem tariffs is unlikely to have increased (Berthelon and Freund (2008)). However, institutional and informational trade barriers may have been disproportionately reduced for within-FTA trade and may have resulted in the intensification of short-distance trade if FTAs have predominantly been signed between regional partners.

A first pass on the data suggests that is indeed the case. We measure the naively-corrected-for-FTA distance elasticity δ_{naive} by estimating (3) in each year while augmenting the vector S_t with a separate control for each active FTA and an additional control for GATT/WTO membership. This method reverses the direction of change in the distance elasticity (fig.5). The effect

[20] Explained variation in the stable sample is 79% with the fixed, and 51% with the current product bundle.

of distance decreases by 1% per year without controlling for sectoral composition effects (left pane), and it decreases by .5% per year when the composition of the country bundle is fixed (right pane)[21].

Figure 5: Evolution of the distance elasticity with FTA controls

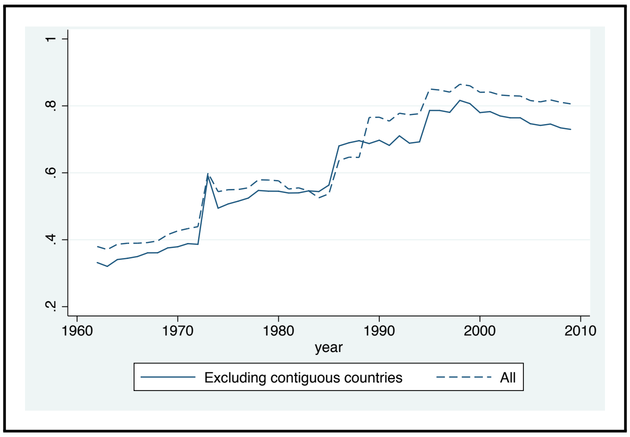


However, this results are not enough to show that the rise of FTA explains the distance puzzle by itself. Fig.6 reports the increase in FTA coverage of trade taking place at less than 2000 km from 40 to 80% in 1963-2009. For the sake of the argument, imagine that FTA have had actually no trade-creating effect and that their coverage has grown to include all nearby-trade in 2009. In that case, an exogenous increase of the distance elasticity manifested by an increase of nearby-trade relative to distant-trade would be completely masked by the introduction of FTA controls.

One way to explore this objection is to verify the endogeneity of FTA formation. If some, deep, FTA might plausibly have trade-creation effects, others, shallow, FTA might not. Bosquet and Boulhol (2009) investigate the impact of FTA formation on the distance puzzle while controlling for the endogeneity of country selection into FTAs with the Baier and Bergstrand (2007) methodology. Constraining the trade creating effect of FTAs to be identical in each year and for all FTAs, Bosquet and Boulhol (2009) (compare their graph 9 and 11) find that FTAs have had no impact on the evolution of the distance elasticity, even though numerous other papers have found that, when controling for endogeneity, there was a trade-creating effect of FTA (Baier and Bergstrand (2007), Egger et al. (2011), Baier et al. (2014)).

[21] App.B provides details on each FTA and the years in which it appears in the sample.

Figure 6: Share of intra-FTA trade among nearby countries (2000km or less)



1.4 Summing up: the robustness of the distance puzzle

Table 1: Evolution of δ_t: sample, composition and FTA effects

	Full sample			Stable sample		
	rate (%)	R-sq	tot.change	rate (%)	R-sq	tot.change
Baseline	.10**	.08	1.049	.44***	.51	1.225
Stable world product bundle	.29***	.56	1.143	.44***	.67	1.224
Stable country product bundle	.43***	.89	1.218	.64***	.79	1.339
FTA	−.88***	.48	0.667	−1.99***	.55	0.398

Note: Estimated annual growth rates reported in col.2 and col.5 are obtained as a geometric fit on the basis of annual point estimates of the distance coefficient in 1963-2009. Col.3 and col.6 report the share of time variation in the point estimate explained with the annualized growth rate.

Table 1 summarizes our findings. The distance puzzle is exacerbated in the sample of stable pairs and robust to fixing the product composition of world trade. Hence, the non-decreasing distance elasticity is likely to be a structural outcome rather than an artefact of composition effects. FTAs might explain the distance puzzle, but there might also be a simple selection effect, e.g. through the increasing geographic scope of FTAs and their increasing coverage of short-distance trade rather than through the intensification of within-FTA trade. These results motivate our focus on structural heterogeneity as a possible alternative explanation of the non-decreasing distance elasticity.

2 Interpreting the distance coefficient

2.1 What does the trade elasticity actually measure?

The distance coefficient is the product of two elasticities: the elasticity of trade costs to distance ρ_t and the elasticity of trade flows to trade costs ζ_t. The question addressed in this paper is very simple. Have trade flows become more sensitive to trade costs (increasing ζ), or have trade costs become more sensitive to distance (increasing ρ)? In section 3 we estimate the evolution of ζ to deduce the evolution of ρ. The three main microfoundations of the gravity model of trade give structurally different interpretations to ζ but not to ρ. In this section, we provide details on the procedure used in this paper to estimate the evolution of ζ in the Armington framework.

In Eaton and Kortum (2002) the heterogeneity dimension captured by the trade elasticity ζ is intersectoral. Consumers maximize a CES utility function defined at the intersectoral level by shopping around the world for the cheapest supplier of each sectoral good.

In the Melitz-Chaney framework the heterogeneity dimension captured by the trade elasticity ζ is intrasectoral. Consumers value firm-specific varieties of sectoral goods which they acquire in monopolistic competition markets.

In Anderson and van Wincoop (2003) there is no heterogeneity in productive efficiency. The production process in each country and sector is constant returns to scale. There is thus perfect competition between domestic producers. The heterogeneity dimension comes from the assumption that consumers perceive products of different national origin as intrinsically imperfect substitutes. Each country is specialized in production of country-specific varieties which on aggregate give a country-specific composite good. The parameter ζ is the lower tier Armington elasticity of substitution. It measures the degree of substitutability of goods of different national origin.[22]

2.2 What do we know about the evolution of the trade elasticity?

In neither of these models there is a theoretical mechanism to explain a change in the trade elasticity overtime. A shock to consumer preferences or to the shape parameter of the productivity distribution would be required. Nonetheless, the heterogeneity parameter measured on

[22] The upper-tier elasticity measures substitutability of domestic products and an aggregate import good. The lower-tier one measures substitutability between importers of a given good. See Sato (1967); Reinert and Shiells (1991); Saito (2004)

aggregate trade data could have evolved over 1962-2009 without any shock to the underlying heterogeneity, either through changes in the range of traded goods, time-sensitive aggregation issues linked to estimating a single parameter across sectors, or agents' adaptation to a changing economic environment. However, there is little empirical evidence on the evolution of sector-specific and aggregate trade elasticites in either model.

To the best of our knowledge, there is no direct empirical evidence on the evolution of sectoral or aggregate efficiency dispersion parameters in the heterogenous firms framework. Different theoretically grounded methods have been used to estimate aggregate trade elasticities in the Ricardian framework for a specific year(Eaton and Kortum (2002); Simonovska and Waugh (2014); Costinot et al. (2012); Caliendo and Parro (2012)), but only Levchenko and Zhang (2011) study the evolution of intersectoral productivity dispersion. They find evidence of within-country convergence in sectoral knowledge stocks in 1960-2010. As there is less heterogeneity in producer efficiency across the set of goods comparative advantage exerts a weaker force against trade resistance imposed by trade barriers.

Evidence on the evolution of lower-tier Armington elasticities of substitution measured on aggregate data is scarce. For France, Welsch (2006) provides estimates of aggregate lower-tier Armington elasticities since the 1970s. He finds that among exporters to the French market the elasticity peaked in the 1970s, and progressively decreased in the 1990s. Broda and Weinstein (2006) provide evidence on the evolution of sectoral Armington lower-tier elasticities between 1972-1988 and 1990-2001 for American imports. They find that they have decreased for all types of goods at all levels of product disaggregation, i.e. at the 10-digit, 5-digit, and 3-digit levels. These results indicate that the parameter estimated on aggregate trade data would also have decreased, deepening the distance puzzle. But to the best of our knowledge, no paper has as yet provided evidence on the evolution of Armington elasticities for aggregate bilateral trade while constraining the parameter to be the same across destination markets.

2.3 A method to measure trade elasticity in the Armington framework

The UN COMTRADE bilateral trade database covers the majority of countries over 1963-2009. It gives information on trade flows and cif unit values at the SITC 4-digit level. This data are sufficient to estimate the trade elasticity in the Armington framework. If we have importer-specific prices in destination markets and importer-specific market share, we can estimate a market share equation at the level of aggregated bilateral trade. A well-known result that in the one-good Armington framework, assuming CES utility functions is::

$$X_{ij} = \left(\frac{P_{ij}}{P_j}\right)^{-(\sigma-1)} Y_j \tag{4}$$

where X_{ij} is the cif value of the exports from i to j, P_{ij} is the cif price of the good shipped from i to j, P_j is the import price index in the destination and Y_j is the demand for imported goods in destination.[23] The exponent $(\sigma - 1)$ captures substitutability of country-composite goods across frameworks. It is also the aggregate trade elasticity ζ in the Armington framework.

Briging this equation to the data is difficult, however, as we do not observe aggregate prices, but unit values at the SITC 4-digit category level. Still, the distance puzzle concerns an elasticity estimated on aggregate trade data. As shown by Imbs and Méjean (2011) this parameter cannot generally be mimicked by a theoretically grounded weighted average of sector-specific trade elasticities. Hence, we need an estimation procedure that works directly with aggregate data.

Define aggregate imports from source i as the sum of imports from each sector k where a sector corresponds to a SITC 4-digit category: $X_{ij} = \sum_k X_{k,ij}$.[24] To preserve our ability to proceed with linear methods of estimation, we will assume that market share equation for aggregate bilateral trade as a function of the weighted average of sectoral relative prices of i in j is:

$$X_{ij} = \left[\sum_{k=1}^{K} \omega_{k,j} \frac{P_{k,ij}}{P_{k,j}}\right]^{(1-\sigma)} Y_j \tag{5}$$

Where $\omega_{k,j} = \frac{Y_{k,j}}{Y_j}$ is the share of sectoral demand in country j in sector k, $P_{k,ij}$ the cif price of the good from sector k shipped from i to j and $P_{k,j}$ the sectoral import price index in destination j and sector k. Using COMTRADE, the unit values of SITC 4-digit categories are interpreted as source-specific sectoral prices $P_{k,ij}$ Destination-specific sectoral import prices $P_{k,j}$ are constructed as a weighted average of observed unit values for each source in sector k where weights are given by the market share of each source in this sector in the destination: $P_{k,j} = \sum w_{k,ij} P_{k,ij}$ with $w_{k,ij} = X_{k,ij}/Y_{k,j}$. Exporters for which some trade but no unit value is observed in k are excluded from the computation of $P_{k,j}$.

In other words, with the assumption that σ is constant accross sectors and countries, the market share equation for aggregate bilateral trade can be written as a function of the weighted

[23] This assumes that the cif price is the consumer price in the importing country. This is not strictly true because custom duties are not included in the cif price. However, origin-specific variations in custom duties are much smaller than variations in cif prices.

[24] When several quantity units are observed, the sector is defined at the product*quantity-unit level.

average of sectoral relative prices of each source in the destination. Each sectoral relative price is weighted according to the share of the sector in total expenditure of the destination which means that identical weights are applied on both sides of the equation.[25]

Transforming equation 5 gives the equation which is estimated in cross-section:

$$\ln(X_{ij}/Y_j) = \lambda_0 - (\sigma - 1)\ln\left(\sum_k \omega_{k,j}\frac{P_{k,ij}}{P_{k,j}}\right) + fe_{exp} + fe_{imp} + \eta_{ij} \tag{6}$$

where fe_{exp} and fe_{imp} are source and destination fixed effects, η_{ij} is a multiplicative error term, and λ_0 is a constant. Source fixed effects control for the world preference for products of this origin. Destination fixed effects control for unobserved domestic prices. The PPML estimator is used because of the heteroskedasticity of the market share equation in levels.

3 Evolution of the Armington trade elasticity in 1963-2009

3.1 The incidence of missing unit values

To estimate the market share equation on the COMTRADE dataset we need to tackle the question of missing information on trade flows and unit values (uv).

A first difficulty arises when the trade flow is observed but information on quantities is missing, and it is therefore not possible to compute the unit value. On average, lacking uv corresponds to 14% of total recorded trade in 1962-2009, with a gradual decrease from 17% to 10% between 1962-2000, and a subsequent increase back to 18% in 2001-2009. In 2001-2006 it is 85-87%, and about 82% in 2007-2009. We assume that information on quantities is missing due to imperfections in the data collection procedure, and that bilateral trade flows are observed with a similar degree of precision whether or not quantity had been recorded. To deal with missing uv, we impute prices from similar products using a stepwise price imputation procedure[26].

The stepwise price imputation procedure is as follows. The relative price of each source in the destination is constructed at the highest disaggregation level for each product and quantity unit in which the source is active, the *4'-digit* level. We then proceed level by level for aggregation: the relative price of the composite sectoral good of the source is constructed at

[25] To see this, rewrite aggregate imports: $\frac{X_{ij}}{Y_j} = \sum_k \frac{X_{k,ij}}{Y_{k,j}}\frac{Y_{k,j}}{Y_{,j}}$ where $\frac{Y_{k,j}}{Y_{,j}} = \omega_{k,j}$.

[26] The weighted average price at a higher aggregation level for this sector and source will be used for the unobserved price. An alternative procedure consists in imputing the relative price observed at the same disaggregation level for another source with a similar market share in this sector and destination. Results are not sensitive to the procedure used.

the 4-digit level using the weighted average relative price observed at the 4'-digit level, with destination-specific weights for each variety of the 4'-digit good the source is active in. Given relative prices constructed at the 4-digit level, destination-specific weights are used to aggregate these up to the 3-digit level, and so on until the relative price for the composite good is constructed using relative prices at the 1-digit level. This improves the estimation of prices if one assumes that missing destination-specific relative prices at the 4'-digit can be approximated by the mean observed destination-specific relative price among the corresponding 4-digit group (and similarly at each aggregation level).

3.2 Zero trade flows

A second difficulty arises when both quantity and trade data are missing. Zero trade flows (ztf) are a prevalent feature of the data while under model assumptions some trade should be observed in every sector k between all pairs ij.[27] We assume that this information is missing because the underlying trade flow is positive but so small that it does not pass the threshold applied by the data collecting authorities (in UN COMTRADE this threshold corresponds to 1000 USD). Such flows, if recorded, would not substantially modify the distribution of observed market shares in the destination (the left hand side of (6)) because they are an order of magnitude smaller than observed trade.

We use the same stepwise price imputation for zero trade flows as in the case of missing unit values. This is problematic because statistically unobserved trade values must correspond to a higher cif price than the maximum observed price in the destination across all sources and sectors while by construction we postulate that unobserved relative prices in ztf sectors are equal to a weighted average relative price across sectors in which trade is observed. [28]

This assumption would not bias our estimate if the underestimation factor were constant across exporters. This scalar would cancel out across sources, and the estimated substitutability parameter would correspond to the true parameter. Table 2 shows it is not the case. The share of ztf is strongly decreasing in market share, i.e. the underestimation factor is larger for small exporters (though they already have higher prices). As a result, for a given observed distribution of market shares, the underlying dispersion in relative prices of the composite good is greater

[27] The share of observed SITC 4-digit flows relatively to the total number of potential SITC 4-digit bilateral trade flows increases from 10% to 14% in 1962-2009.

[28] An alternative method consists in imputing unobserved relative prices with some arbitrary price above the maximum observed in the destination. As ztf constitute 85-90% of all 4-digit trade flows, this method is problematic because results are driven by imputed rather than effectively observed prices.

than the observed dispersion in relative prices. This means that the estimated parameter $\widetilde{\sigma}$ overestimates the true substitutability parameter σ.

Table 2 shows that the reduction in the share of ztf proceeds at quicker pace in 1962-2009 for small exporters: the coefficient for the interaction term for the market share and year is significant and positive. Table 3 presents the predicted share of ztf for four types of exporters in 1962 and 2009. For a very small exporter with .02% market share, the initial share of ztf is predicted to be .95, and it is reduced to .83 by 2009, i.e. a 12 percentage point decrease. Consider a relatively big exporter, with a 10% market share: its share of ztf is reduced from .72 to .65, a 7 percentage point decrease. As the gap between the share of ztf for big and small exporters is reduced overtime, the overestimation bias of $\widetilde{\sigma}$ is progressively reduced.

Table 2: **Proportion of zero trade flows as a function of market share**

depvar: Share of ZTF				
	(1)	(2)	(3)	(4)
ms	-0.0401*** (0.0001)	-0.2446*** (0.0134)	-0.0427*** (0.0001)	-0.2573*** (0.013)
year	-0.0029*** (0.0000)	-0.0020*** (0.0001)	-0.0033*** (0.0000)	-0.0024*** (0.000)
$ms * year$		0.0001*** (0.0000)		0.0001*** (0.000)
constant	5.3474*** (0.0335)	3.5852*** (0.1372)	6.0976*** (0.0366)	4.2515*** (0.134)
Destination FE	NO	NO	YES	YES
Observations	657001	657001	657001	657001

Notes: The share of ZTF is computed at the SITC 4-digit level. The estimation is conducted in PPML in order to include observations where ztf=0. The log of the market share is used in the estimation. Destination fixed effects are included in (3) and (4). Robust standard errors are in parentheses. *** p<0.01.

Table 3: **Predicted share of ztf for exporters with different market share, 4-digit level**

year	ms=0.02%	ms=1%	ms=10%	ms=28.7%
1962	0.95	0.80	0.72	0.69
2009	0.83	0.71	0.65	0.62

Columns (1) and (4) correspond to the mean and to 2 st. deviations above the mean in the distribution of log market share. Columns (2) and (3) correspond to the mean and to 2 st. deviations above the mean in the distribution of market share.

Thus, the hypothesis we make on unobserved sectoral prices in ztf sectors does not always impede interpreting the evolution of the underlying substitutability parameter. In particular, because the overestimation bias is reduced overtime, if it is found that the estimated parameter increases in absolute value, this evolution necessarily provides a lower bound on the increase in the underlying substitutability parameter.

Figure 7 presents the results on the evolution of $(1-\widetilde{\sigma})$ obtained when (6) is estimated on annual crossections of the COMTRADE dataset. The absolute value of trade elasticity has increased by 33% from 1962 to 2009. This corresponds to an annual increase of .6% per year.[29] According the preceeding discussion, this is a lower bound on the increase in the underlying substituability parameter.

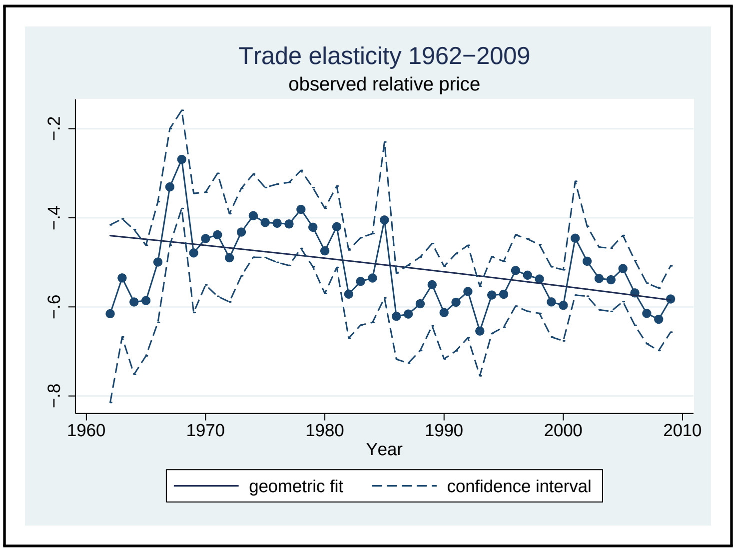


Figure 7: Estimated $(1-\widetilde{\sigma})$

[29] The coefficient of the geometric fit is significant at 1% level.

3.3 Robustness checks

3.3.1 Changing the dataset

We provide a robustness check by estimating the evolution of the heterogeneity parameter for aggregate bilateral trade on a different dataset. We use the BACI dataset which reports bilateral trade data at the HS-1992 6-digit disaggregation level for 1995-2009. The accuracy of the relative prices of country-composite goods constructed with this dataset is improved because the harmonization procedure applied by Gaulier and Zignago (2010) in constructing BACI yields much better-quality unit values while substantially reducing the number of observations with lacking unit value. As a result, at the 6-digit level, less than 7% of total reported trade in BACI has missing unit values. This is reduced to 1-3% of total trade when the data is aggregated to the 4-digit level, as opposed to more than 10% in the raw COMTRADE data we originally used. Another advantage is that the share of ztf in BACI is stable in 1995-2009 as opposed to relatively strong fluctuations in the share of ztf overtime in our original dataset. The disadvantage of BACI is that it covers only a relatively short period compared to the years over which the distance puzzle exists. Obviously, we do not expect to reproduce exactly the results obtained with our original dataset because the trade classification and its level of aggregation are different.

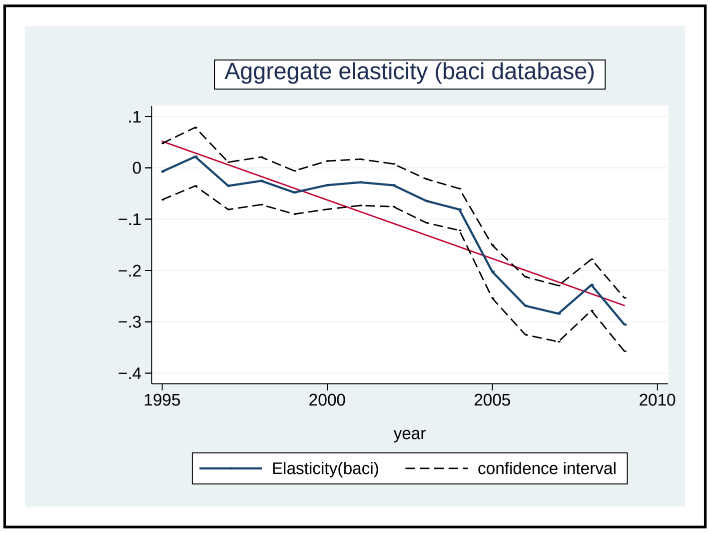


Figure 8: Estimated $(1-\widetilde{\sigma})$, BACI database

Fig. 8 shows that our results hold: the elasticity parameter is found to increase in absolute value from 1995-2009. This can be compared with the equivalent period in our original dataset: the increase in the elasticity is much steeper on the BACI dataset. This finding supports the idea that our benchmark estimation likely provides a lower bound on the increase in the aggregate trade elasticity. However, the level of the elasticity estimated in 1995-1999 on BACI data is puzzling and suggests the existence of an attenuation bias. This is the focus of our second robustness check.

3.3.2 Instrumenting: motivation and results

The results just presented are subject to caution if supply schedules are not horizontal.[30] The demand elasticity parameter estimated in the market share equation would then be subject to attenuation bias due to not controlling for potentially positive and finite supply elasticities. This attenuation bias would not be problematic for analyzing the evolution of the substitutability parameter if only the level of the parameter were affected. The problem arises because, as shown by Feenstra (1994), the attenuation bias also impacts the evolution of the parameter.

Feenstra has developed an instrumental variables' approach which solves this problem (see Feenstra (1994), and refinements in Broda and Weinstein (2006) and Imbs and Méjean (2011)). This method exploits year-to-year variations in relative prices and market shares over 10- to 20-year estimation windows to compute the Armington elasticity. We do not use this approach for two sets of reasons. First, Feenstra's method relies on the assumption that the elasticity parameter remains constant through time, whereas we allow the parameter to vary in each year. Second, more fundamentally, Feenstra's elasticity parameter determines short-run, marginal, longitudinal effects whereas we are interested in the elasticity parameter which determines long-run, equilibrium, cross-section outcomes. It is not immediate that these two elasticities should be the same.

We adopt a different approach that preserves the time dimension which is central to our analysis. We need an instrument which adequately captures exporter-specific shocks to the price of the composite good which are not demand-driven, such as exogenous shocks to inputs' prices. We would like to use changes in the bilateral-specific real exchange rate. One possibility would be to use Producer Price Index (PPI) since it captures the evolution of prices faced by producers on the inputs' side. Unfortunately, we do not have PPI data for most countries and years in our sample. We therefore settle for an alternative exporter-specific price level indicator: the GDP price level in current US dollars as reported in the Penn World Tables for 189 countries

[30] Broda et al. (2008) find that supply elasticities are finite at the 4-digit level. On the other hand, Magee and Magee (2008) find that the small country assumption may hold in the data in which case there would be no attenuation bias.

in 1950-2009.[31]

The instrumenting procedure is the following. First, we compute relative prices for exporter-specific composite goods in each destination market using the stepwise price imputation procedure (see 3.1). Second, for each destination market, we compute the mean evolution of GDP price levels in current US dollars of its trading partners, weighted by their market shares in this destination. This amounts to computing the evolution of the relevant real exchange rate for each specific bilateral trade relation. Third, we compute a hypothetical relative price at time t for each exporter in each market as the product of its relative price at time $(t-1)$ and the evolution of its GDP price level between t and $(t-1)$ relatively to all other trading partners in this destination. Fourth, we predict the relative price of each exporter in each destination at time t by regressing its observed relative price on this hypothetical relative price. This gives an instrumented relative price for each exporter which depends only on its past relative price and the relative evolution of its GDP price level. Finally, we estimate (6) using these instrumented relative prices instead of the observed relative prices.

It could be argued that allowing for just one lag inadequately captures the temporal relationship between shocks to inputs' prices and their pass-through to the price of exported output. Indeed, if prices are relatively persistent, the instrumenting procedure would amount to little more than replacing observed prices in t with lagged observed prices in $(t-1)$. We therefore also estimate (6) using as instrument the evolution of each exporter's GDP price level relatively to all other trading partners in the destination between $(t-s)$ and t where $s = 1, ..., 10$.

Results obtained with one lag ($s = 1$) are shown in Fig.9. The absolute value of the substitutability parameter has increased by 13% in 1963-2009 while the level of the estimated parameter increases by 9% on average relatively to the estimate obtained with non-instrumented prices.

This result is robust to increasing the number of periods in which the evolution of exports' prices is predicted with the evolution of domestic prices. Thus, in 1972-2009, the elasticity increases by 20% when the instrument is constructed with one lag, and by 23% when the number of lags is 10 (see Appendix C). The evolution of the parameter becomes steeper as we increase the number of lags. Therefore, it is likely that our estimate provides a lower bound on the increase in the true substitutability parameter.

[31] See Heston et al. (2011). Our sample of countries does not coincide perfectly with the countries in the Penn World Tables. However, as it is the smallest exporters which drop out, the sample adjustment in terms of world trade coverage is minor.

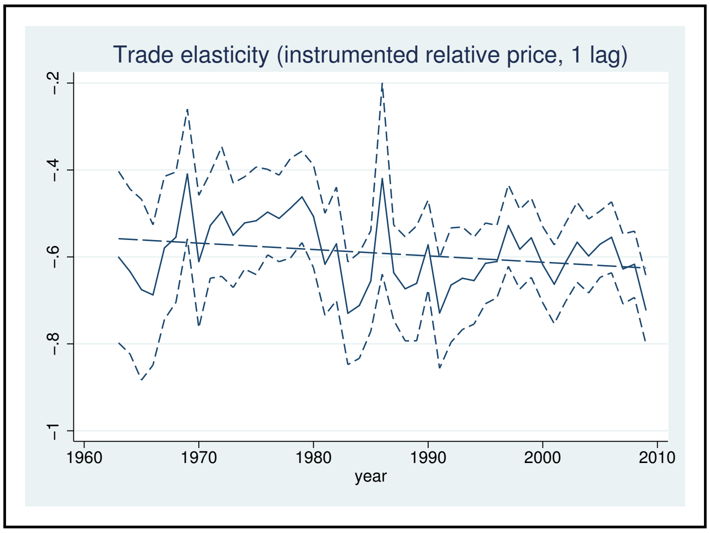


Figure 9: Estimated $(1-\widetilde{\sigma})$, instrumented relative price of composite good, 1 lag

3.4 Is there a distance puzzle left?

This section has provided empirical evidence on the evolution of the aggregate substitutability parameter for world trade in 1963-2009. This substitutability parameter corresponds to the aggregate trade elasticity in the Armington framework. We find that this parameter has increased by 33% between 1963 and 2009 in the benchmark estimation, and by 13% when prices are instrumented. Both estimates are likely to be lower bounds on the increase in the true substitutability parameter. Section 1 has shown that the distance elasticity of trade has increased by 7% over the same period. Combining these two results, there is no distance puzzle in the framework of the Armington model in as much as the elasticity of trade costs to distance has decreased by at least 5-7% in 1963-2009.[32] Increasing perceived substitutability of country-specific composite goods contributes to the increasing distance elasticity of trade.

The reduction in the elasticity of trade costs to distance is even more pronounced if we focus on 1970-2009. As shown by Hummels (2007), this period is characterized by a new phenomenon: the fact that air transportation starts playing a substantial role in world trade. The instrumented Armington elasticity increases by 19% in this period while the evolution of the distance elasticity is best described as flat. It follows that the elasticity of trade costs to distance has decreased by at least 17% in 1970-2009.

What is the economic interpretation of an increasing substitutability parameter measured on aggregate data?

[32] The elasticity decreases by 5% when the evolution of ρ is computed from the ratio of trends, and by 7% when it is computed as the trend of the ratios.

First, the degree of perceived similarity of country-composite goods may have increased. Since the 1960s, a growing number of countries started producing a set of goods similar to that of developed countries. This process has increased the number of available varieties and, potentially, their degree of similarity.[33]

Second, composition effects may have lead to changes in the parameter estimated on aggregate data. If the reduction in trade barriers has led to the expansion in the range of traded goods, trade in previously non-traded sectors could modify measured substitutability of country-composite goods. Non-uniform reductions in sectoral trade costs would also modify the composition of world trade, leading to a change in the substitutability parameter measured on aggregate data. However, at first approximation, the rising importance of manufactures compared to primary products in world trade should have reduced substituability.

Ideally, we would like to separate out the impact of composition and sector-specific effects to quantify the net effect of increased perceived similarity of country-composite goods. This is however impossible because, as shown by Imbs and Méjean (2011), the parameter estimated on aggregate data cannot be mimicked by a weighted average of sectoral parameters. The bottom line is that an increase in measured substitutability for country-composite goods is consistent with complex competition dynamics in price and quality documented by Amiti and Khandelwal (2013) as well as with increased vertical specialization of countries within sectors documented by Fontagné et al. (2008).

Conclusion

The estimated effect of distance in gravity equations has increased in the past fifty years despite substantial innovation in transportation and communication: this is the 'distance puzzle'. Using COMTRADE 4-digit bilateral trade data in 1963-2009, this paper finds that the evolution of the elasticity of trade flows to trade costs, referred to as the 'trade elasticity', provides a direct explanation of the increasing distance elasticity of trade. Increased sensitivity of trade flows to relative prices has more than compensated the reduction in the elasticity of trade costs to distance.

The paper proceeds in three steps. First, it shows that the distance puzzle is a feature of our data by estimating yearly cross-section gravity equations. In the baseline estimation the

[33] Schott (2004) documents increased similarity in the set of exported goods of US trade partners while Broda et al. (2006) document the increase in the number of imported varieties since the 1970s.

distance coefficient has increased by 7% from 1962 to 2009. This result holds when we correct for changes in the sample of trading partners and the composition of world trade. Taking into account FTAs seems to solve the distance puzzle, but this might be an artefact of their growing importance: introducing FTAs dummies amounts to adding a time-growing number of proximity controls in the estimation.

Second, the paper suggests a straigthforward method of measuring structural heterogeneity in the Armington framework. In the main theoretical foundations of the gravity equation the distance coefficient is the product of the elasticity of trade costs to distance and a measure of heterogeneity. In the Armington framework, heterogeneity is inter-country and intra-sector. The trade elasticity corresponds to the degree of perceived substitutability of country-specific varieties of each good, which can be approximated by studying the relations between the price level and the market share of importers in destination markets.

Third, the paper estimates the evolution of the trade elasticity in the Armington framework, i.e. the substitution elasticity between country composite goods. It uses 4-digit unit values as proxies for sectoral prices. In our method, unobserved unit values for zero trade flows lead to an overestimation bias that is reduced over time. As the estimated elasticity still increases in absolute value this evolution provides a lower bound on the increase in the absolute value of the underlying trade elasticity. Once instrumented by bilateral real exchange rates, the estimated elasticity increases by 13% between 1963 and 2009. The evolution of the distance coefficient is thus compatible with a decrease of the elasticity of trade costs to distance of at least 5 to 7%. This reduction in the elasticity of trade costs to distance is even more pronounced if we focus on the period in which air transportation starts playing an important role in bilateral trade. We find that from 1970 to 2009 the elasticity of trade costs to distance has decreased by 17% while the perceived substitutability of countries' product bundles has increased by at least 19%.

References

Amiti, M. and A. Khandelwal (2013). Import competition and quality upgrading. *Review of Economics and Statistics 95*(2), 476–490.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Anderson, J. E. and E. v. Wincoop (2004). Trade costs. *Journal of Economic Literature 42*(3), 691–751.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Baier, S. L. and J. H. Bergstrand (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics 71*(1), 72–95.

Baier, S. L., J. H. Bergstrand, and M. Feng (2014). Economic integration agreements and the margins of international trade. *Journal of International Economics In Press*.

Baldwin, R. and J. Harrigan (2011). Zeros, quality and space: trade theory and trade evidence. *American Economic Journal: Microeconomics 3*(2), 60–88.

Berthelon, M. and C. Freund (2008). On the conservation of distance in international trade. *Journal of International Economics 75*(2), 310–320.

Bosquet, C. and H. Boulhol (2009). Gravity, log of gravity and the distance puzzle. *GREQAM Working Paper 2009*(12).

Bosquet, C. and H. Boulhol (2010). Scale dependence of the negative binomial pseudo-maximum likelihood estimator. *CES Working Papers 2010*(92).

Broda, C., J. Greenfield, and D. Weinstein (2006). From groundnuts to globalization: A structural estimate of trade and growth. *NBER Working Paper Series 12512*.

Broda, C., N. Limão, and D. Weinstein (2008). Optimal tariffs: The evidence. *American Economic Review 98*(5), 2032–2065.

Broda, C. and D. E. Weinstein (2006). Globalization and the gains from variety. *Quarterly Journal of Economics 121*(2), 541–585.

Brun, J.-F., C. Carrère, P. Guillaumont, and J. d. Melo (2005). Has distance died? evidence from a panel gravity model. *World Bank Economic Review 19*, 99–120.

Buch, C. M., J. Kleinert, and F. Toubal (2004). The distance puzzle: on the interpretation of the distance coefficient in gravity equations. *Economics Letters 83*, 293–298.

Cairncross, F. (1997). *The Death of Distance: How the Communications Revolution will change our Lives*. London: Orion Business Books.

Caliendo, L. and F. Parro (2012). Estimates of the trade and welfare effects of nafta. *NBER Working Paper Series* (18508).

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–1721.

Chaney, T. (2013). The gravity equation in international trade: an explanation. *NBER Working Paper Series 19285*.

Combes, P.-P., T. Mayer, and J.-F. Thisse (2006). *Economie Géographique*. Paris: Economica.

Costinot, A., D. Donaldson, and I. Komunjer (2012). What goods do countries trade? A quantitative exploration of Ricardo's ideas. *Review of Economics and Statistics (forthcoming)*.

Crawford, J.-A. and R. V. Fiorentino (2005). The changing landscape of regional trade agreements. *WTO Discussion Papers* (8).

Daudin, G. (2003). La logistique de la mondialisation. *Revue de l'OFCE* (87), 411–435.

Daudin, G. (2005). Les transactions de la mondialisation. *Revue de l'OFCE* (92), 223–262.

Daudin, G., C. Rifflart, and D. Schweisguth (2011). Who produces for whom in the world economy. *Canadian Journal of Economics 44*(4), 1403–1437.

Disdier, A.-C. and K. Head (2008). The puzzling persistence of the distance effect in bilateral trade. *The Review of Economics and Statistics 90*(1), 37–48.

Duranton, G. and M. Storper (2008). Rising trade costs. agglomeration and trade with endogenous transaction costs. *Canadian Journal of Economics 41*(1), 292–319.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Egger, P., M. Larch, K. E. Staub, and R. Winkelmann (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy 3*(3), 113–143.

Erkel-Rousse, H. and D. Mirza (2002). Import price elasticities: reconsidering the evidence. *The Canadian Journal of Economics 35*(2).

Fally, T. (2012). Structural gravity and fixed effects. *mimeo*.

Feenstra, R. (1994). New product varieties and the measurement of international prices. *The American Economic Review 84*(1), 157–177.

Felbermayr, G. J. and W. Kohler (2006). Exploring the intensive and extensive margins of trade. *Review of World Economics 142*(4), 642–674.

Fontagné, L., G. Gaulier, and S. Zignago (2008). North-south competition in quality. *Economic Policy 23*(53), 51–91.

Fontagné, L. and S. Zignago (2007). A re-evaluation of the impact of regional trade agreements on trade patterns. *Economie Internationale 109*, 31–51.

Friedman, T. L. (2007). *The World is Flat: A Brief History of the Twenty-first Century, 2nd edition*. Farrar Straus & Giroux.

Gaulier, G. and S. Zignago (2010). Baci: International trade database at the product level. *CEPII working papers* (23).

Head, K. and T. Mayer (2013). What separates us? sources of resistance to globalization. *Canadian Journal of Economics 46*(4), 1196–1231.

Head, K. and T. Mayer (2014). *Handbook of international economics*, Volume 4, Chapter Gravity Equations: workhorse, toolkit, and cookbook. Elsevier.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: trading partners and trading volumes. *Quarterly Journal of Economics 123*(2), 441–487.

Heston, A., R. Summers, and B. Aten (2011). *Penn world table version 7.0.* Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hummels, D. (2007). Transportation costs and international trade in the second era of globalization. *Journal of Economic Perspectives 21*(3), 131–154.

Hummels, D. and G. Schaur (2013). Time as a trade barrier. *American Economic Review 103*, 1–27.

Imbs, J. and I. Méjean (2011). Elasticity optimism. *CEPR Discussion Paper 7177*.

Krautheim, S. (2012). Heterogeneous firms, exporter networks, and the effect of distance on international trade. *Journal of International Economics 87*, 27–35.

Larch, M., P.-J. Norbäck, and S. Sirries (2013). Heterogenous firms, globalization, and the distance puzzle. *IFN Working Paper* (957).

Levchenko, A. A. and J. Zhang (2011). The evolution of comparative advantage: Measurement and welfare implications. *NBER Working Paper Series 16806*.

Levinson, M. (2006). *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton: Princeton University Press.

Magee, C. S. and S. P. Magee (2008). The united states is a small country in world trade. *Review of International Economics 16*(5), 990–1004.

Manning, W. G. and J. Mullahy (1999). Estimating log models: To transform or not to transform? *NBER Technical Working Paper Series 246*.

Mayer, T. and S. Zignago (2011). Notes on CEPII's distances measures: the GeoDist database. *MPRA Working Paper Series* (36347).

Noguera, G. (2012). Trade costs and gravity for value added trade. *unpublished manuscript*.

Reinert, K. A. and C. R. Shiells (1991). Trade substitution elasticities for analysis of a north american free trade area. *Staff Research Study, USITC* (19).

Saito, M. (2004). Armington elasticities in intermediate inputs' trade: a problem in using multilateral data. *The Canadian Journal of Economics 37*(4), 1097–1117.

Santos Silva, J. M. C. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Santos Silva, J. M. C. and S. Tenreyro (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters 112*, 220–222.

Sato, K. (1967). A two-level constant-elasticity-of-substitution production function. *Review of Economic Studies 34*, 201–218.

Schott, P. (2004). Across product vs. within-product specialization in international trade. *The Quarterly Journal of Economics 119*(2), 647–678.

Simonovska, I. and M. Waugh (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics 92*(1).

Welsch, H. (2006). Armington elasticities and induced intra-industry specialization: The case of france, 1970-1997. *Economic Modelling 23*, 556–567. Armington elasticities' estimates for France '70-'90s.

Yi, K.-M. (2010). Can multistage production explain the home-bias in trade? *American Economic Review 100*(1), 364–393.

A Full and superbalanced samples

The full sample contains 207 reporters (R) and 230 partners (P) which are listed in tables 4 and 5 below. ‘S’ indicates that the country is present in the superbalanced sample.

In the full sample, several countries shift from reporting trade on an individual basis to reporting trade jointly with another country. This is the case of Belgium and Luxembourg, as well as Eritrea and Ethiopia. For consistency, we use a single country identifier for each of these two pairs. A single country identifier is also used for Yougoslavia and for Serbia and Montenegro. Fig.10 shows the distribution of pairs in the full sample according to the number of years in which the pair reports a positive amount of trade.

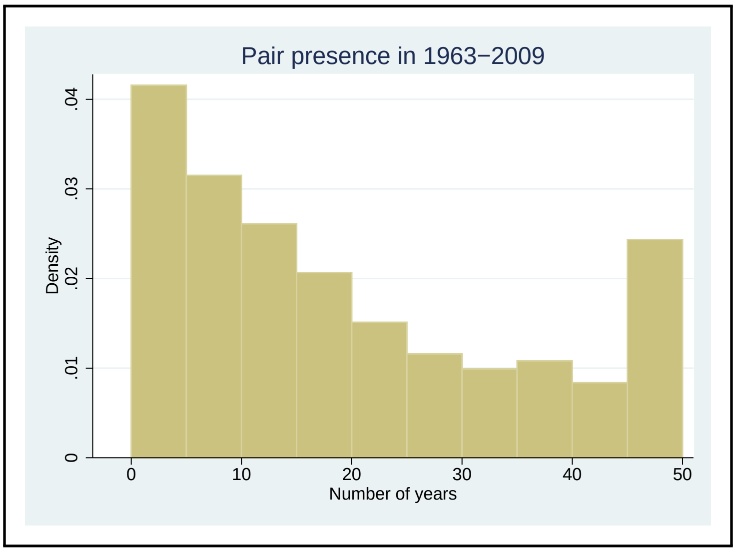


Figure 10: Number of years each pair is present in the sample

The superbalanced sample corresponds to the subsample of pairs which trade both ways in each and every year in 1963-2009. To avoid discarding pairs which fall out of the superbalanced sample because countries split up or reunite at some point in 1963-2009, we introduce several additional single country identifiers before constructing the superbalanced sample. Consequently, Germany is present in the superbalanced sample.[34]

The superbalanced sample comprises 1056 trading pairs and corresponds to 37 countries.This is less than the 1332 pairs which would be observed if each reporter traded both ways with every

[34] A single identifier is used for for East, West, and reunited Germany. Similarly, we use a single identifier for the Czech Republic, Slovakia, and Czechoslovakia; and another single identifier for the USSR and the 15 countries which were formed after USSR split up. The 15 countries which constituted the USSR are absent from the superbalanced sample because the USSR is never a reporter to COMTRADE. The Czech Republic and Slovakia also drop out because they do not have reciprocal trade in all years with another country of the sample.

other country. Indeed, the set of countries which trade with every other country in each year and which we refer to as the 'square sample' comprises just 23 countries (506 pairs). Trade coverage in the superbalanced sample decreases from 70 to 50% of total trade while it is reduced from 60 to 40% for the square sample (see Fig.11).

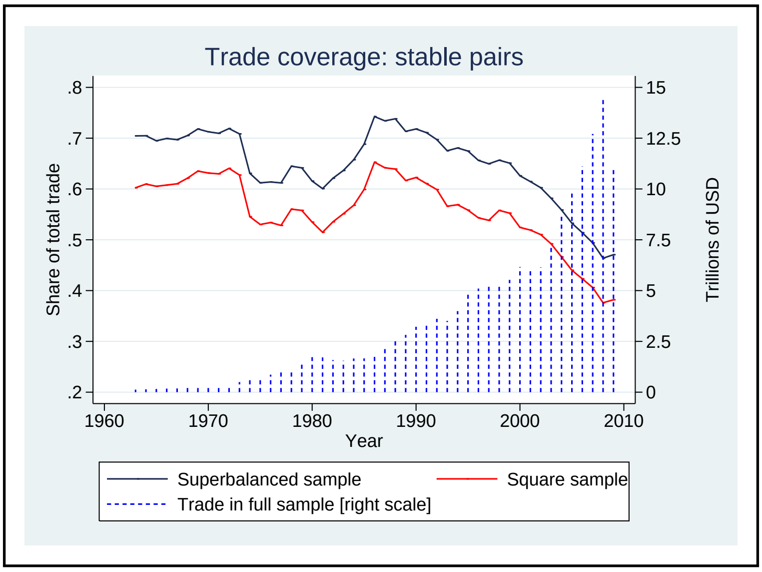


Figure 11: Trade coverage in 1963-2009

As a check on the year chosen to construct the superbalanced sample we redefine the set of stable pairs as partners trading both ways in each year in 1970-2009. We take 1970 as a comparison benchmark because this is the first year in which the number of reporters in COMTRADE exceeds 100.[35] We obtain a sample of 1604 pairs which correspond to 47 countries (out of 2162 possible pairs).

As illustrated in fig.12, the evolution of trade coverage is not sensitive to the choice of the starting year for constructing the superbalanced sample. Indeed, the main reason for the reduction in trade coverage since the mid-1990s is due to the absence of China and of Central and Eastern European countries from the superbalanced sample. These countries drop out because they do not report trade to COMTRADE until the more recent period.

Another way to check sensitivity to the choice of the starting year is to compute the evolution of trade coverage for the sample of pairs which trade both ways in a specific year.[36] The share of total annual trade attributable to two-way pairs in 1963 and in 1970 is shown in fig.13. As previously, only the level of trade coverage is affected by the choice of the starting year.

[35] There are 80 reporters in 1963, 112 in 1970, 153 in 2000.

[36] The difference with the superbalanced sample is that we relax the constraint that the pair have two-way trade in every year in 1963-2009.

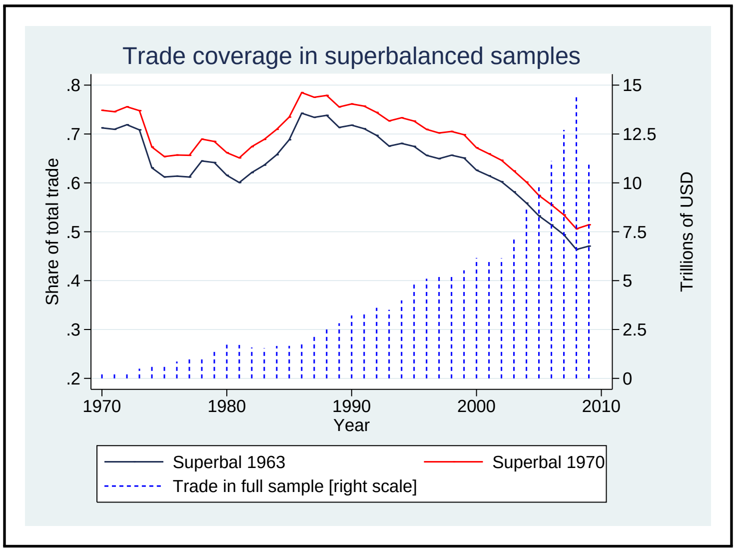


Figure 12: Trade coverage in two superbalanced samples

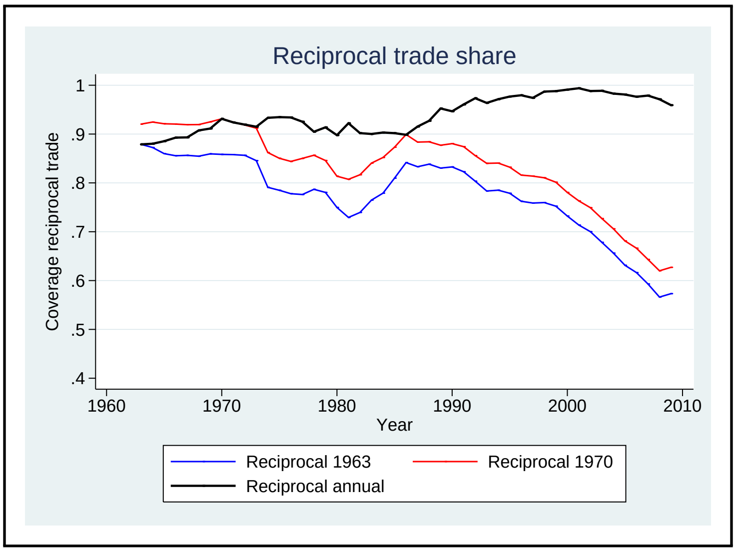


Figure 13: Trade coverage in reciprocal trade

Helpman et al. (2008) have shown that the enlargement of the set of trading partners did not contribute in a major way to the growth of world trade in 1970-1997, with most of the increase being driven by pairs which trade both ways in 1970. We nuance this finding by showing that country pairs which did not have two-way trade in 1970 contributed strongly to the increase in world trade since the mid-1990s. Indeed, while more than 80% (resp. 85%) of world trade in 1990 is attributable to pairs which traded both ways in 1963 (resp. 1970), less than 60% (resp. 70%) is still attributable to such pairs in 2009.

Furthermore, these new trade relationships were formed between countries trading both ways. To illustrate, we graph the share of total trade attributable to pairs which trade both ways in a given year (in black). One-way trade flows represent a marginal and decreasing share of world trade. Since the 1990s, more than 95% of total annual trade takes place between pairs which trade both ways.

The greater frequency of one-way trade relationships observed in the early years of the sample may thus be due to the fact that the number of countries reporting trade to COMTRADE was relatively small. Consequently, reported zeros are likely to be at least in part statistical zeros, i.e. non-zero trade flows reported as zeros due to missing reports to COMTRADE. This evidence also suggests that fixed trade costs are likely to have a bilateral component as in Chaney (2008).

Table 4: List of countries in the full and superbalanced samples

Country name	Status	Country name	Status	Country name	Status
Afghanistan	*R;P*	French Polynesia	*R;P*	N. Mariana Islands	P
Albania	*R;P*	French S. Antartic terr.	P	Norway	*R;P*
Algeria	*R;P*	Gabon	*R;P*	Oman	*R;P*
Andorra	*R;P*	Gambia	*R;P*	Pakistan	*R;P*
Angola	*R;P*	Georgia	*R;P*	Palau	P
Anguilla	*R;P*	**Germany**	**R;P;S**	Panama	*R;P*
Antigua-Barbuda	*R;P*	Ghana	*R;P*	Papua New Guinea	*R;P*
Argentina	**R;P;S**	Gibraltar	P	**Paraguay**	**R;P;S**
Armenia	*R;P*	**Greece**	**R;P;S**	Peru	*R;P*
Aruba	*R;P*	Greenland	*R;P*	**Philippines**	**R;P;S**
Australia	**R;P;S**	Grenada	*R;P*	Pitcairn	P
Austria	**R;P;S**	Guadeloupe	*R;P*	Poland	*R;P*
Azerbaijan	*R;P*	Guatemala	*R;P*	**Portugal**	**R;P;S**
Bahamas	*R;P*	Guinea	*R;P*	Qatar	*R;P*
Bahrain	*R;P*	Guinea-Bissau	*R;P*	Reunion	*R;P*
Bangladesh	*R;P*	Guyana	*R;P*	Romania	*R;P*
Barbados	*R;P*	Haiti	*R;P*	Russian Federation	*R;P*
Belarus	*R;P*	Honduras	*R;P*	Rwanda	*R;P*
Belgium	**R;P;S**	**Hong Kong**	**R;P;S**	St. Helena	P
Belize	*R;P*	Hungary	*R;P*	St. Kitts and Nevis	*R;P*
Benin	*R;P*	**Iceland**	**R;P;S**	St. Lucia	*R;P*
Bermuda	*R;P*	India	*R;P*	St. Vincent-Grenadines	*R;P*
Bhutan	*R;P*	Indonesia	*R;P*	Samoa	*R;P*
Bolivia	*R;P*	Iran	*R;P*	San Marino	P
Bosnia-Herzeg.	*R;P*	Iraq	*R;P*	Sao Tome-Principe	*R;P*
Botswana	*R;P*	Ireland	**R;P;S**	Saudi Arabia	*R;P*
Brazil	**R;P;S**	**Israel**	**R;P;S**	Senegal	*R;P*
Br. Virgin Islands	P	**Italy**	**R;P;S**	Serbia-Montenegro	*R;P*
Brunei Darussalam	*R;P*	Jamaica	*R;P*	Seychelles	*R;P*
Bulgaria	*R;P*	**Japan**	**R;P;S**	Sierra Leone	*R;P*
Burkina Faso	*R;P*	Jordan	*R;P*	**Singapore**	**R;P;S**
Burma	*R;P*	Kazakstan	*R;P*	Slovakia	*R;P*
Burundi	*R;P*	Kenya	*R;P*	Slovenia	*R;P*
Cambodia	*R;P*	Kiribati	*R;P*	Solomon Islands	*R;P*
Cameroon	*R;P*	**Korea**	**R;P;S**	Somalia	*R;P*
Canada	**R;P;S**	DPR of Korea	P	South Africa	*R;P*

Table 5: List of countries in the full and superbalanced samples: Contd.

Country name	Status	Country name	Status	Country name	Status
Cape Verde	*R;P*	Kuwait	*R;P*	Soviet Union	P
Cayman Islands	P	Kyrgyzstan	*R;P*	**Spain**	**R;P;S**
C.African Republic	*R;P*	Lao PDR	*R;P*	Sri Lanka	*R;P*
Chad	*R;P*	Latvia	*R;P*	St. Pierre and Miquelon	*R;P*
Chile	**R;P;S**	Lebanon	*R;P*	Sudan	*R;P*
China	*R;P*	Lesotho	*R;P*	Suriname	*R;P*
Christmas Island	P	Liberia	*R;P*	Swaziland	*R;P*
Cocos Islands	P	Libya	*R;P*	**Sweden**	**R;P;S**
Colombia	**R;P;S**	Lithuania	*R;P*	**Switzerland**	**R;P;S**
Comoros	*R;P*	**Luxembourg**	**R;P;S**	Syria	*R;P*
Congo	*R;P*	Macau (Aomen)	*R;P*	Taiwan	*R;P*
Dem. Rep. of Congo	*R;P*	Macedonia	*R;P*	Tajikistan	*R;P*
Cook Islands	*R;P*	Madagascar	*R;P*	Tanzania	*R;P*
Costa Rica	*R;P*	Malawi	*R;P*	**Thailand**	**R;P;S**
Croatia	*R;P*	**Malaysia**	**R;P;S**	Togo	*R;P*
Cuba	*R;P*	Maldives	*R;P*	Tokelau	P
Cyprus	*R;P*	Mali	*R;P*	Tonga	*R;P*
Czech Republic	*R;P*	Malta	*R;P*	Trinidad-Tobago	*R;P*
Czechoslovakia	*R;P*	Marshall Islands	P	**Tunisia**	**R;P;S**
Côte d'Ivoire	*R;P*	Martinique	*R;P*	**Turkey**	**R;P;S**
Denmark	**R;P;S**	Mauritania	*R;P*	Turkmenistan	*R;P*
Djibouti	*R;P*	Mauritius	*R;P*	Turks-Caicos Islands	*R;P*
Dominica	*R;P*	**Mexico**	**R;P;S**	Tuvalu	*R;P*
Dominican Republic	*R;P*	Micronesia	P	Uganda	*R;P*
East Germany	**R;P;S**	Moldova	*R;P*	Ukraine	*R;P*
East Timor	*R;P*	Mongolia	*R;P*	United Arab Emirates	*R;P*
Ecuador	*R;P*	Montserrat	*R;P*	**United Kingdom**	**R;P;S**
Egypt	*R;P*	Morocco	*R;P*	**USA**	**R;P;S**
El Salvador	**R;P;S**	Mozambique	*R;P*	Uruguay	*R;P*
Equatorial Guinea	P	Namibia	*R;P*	Uzbekistan	P
Eritrea	*R;P*	Nauru	P	Vanuatu	*R;P*
Estonia	*R;P*	Nepal	*R;P*	**Venezuela**	**R;P;S**
Ethiopia	*R;P*	Netherland Antilles	*R;P*	Vietnam	*R;P*
Falkland Islands	P	**Netherlands**	**R;P;S**	Wallis-Futuna	*R;P*
Faroe Islands	*R;P*	New Caledonia	*R;P*	**West Germany**	**R;P;S**
Fiji	*R;P*	New Zealand	*R;P*	Western Sahara	P
Finland	**R;P;S**	Nicaragua	*R;P*	Yemen	*R;P*
Fm Vietnam DR	*R;P*	Niger	*R;P*	Yugoslavia	*R;P*
Fm Vietnam Rp	*R;P*	Nigeria	*R;P*	Zambia	*R;P*
France	**R;P;S**	Niue	P	Zimbabwe	*R;P*
French Guiana	*R;P*	Norfolk Island	P		

B List of included FTAs

Sources: Data for membership of GATT/WTO was taken from the WTO website[37]. Data for common FTA membership was constructed on the basis of Crawford and Fiorentino (2005), Fontagné and Zignago (2007), and the RTA Information System on the WTO website[38].

In squared brackets: *[years in which the FTA appears in the database, where 'F' stands for 'full' and 'S' for 'superbalanced' sample]*

N.B.: The GATT/WTO membership variable is present in the database in all years.

EC (European Communities), then **EU** (European Union): *[1962-2009 (F,S)]*

EFTA (European Free Trade Association): *[1962-2009 (F,S)]*

CACM (Central American Common Market): *[1963-69 and 1993-2009 (F)]*

COMECON (Union of Mutual Economic Assistance): *[1964-1990 (F)]*

CEMAC (Economic and Monetary Community of Central Africa): *[1964-2009, except 1981, 1988, 1991 and 1992 (F)]*

OCT (EC FTA with Overseas Countries and Territories): *[1971-2009 (F)]*

CARICOM (Caribbean Community and Common Market): *[1973-2009 (F)]*

EEA (European Economic Area: EC-EFTA FTA): *[1973-2009 (F,S)]*

PATCRA (Agreement on Trade and Commercial Relations between the Government of Australia and the Government of Papua New Guinea): *[1977-2009 (F)]*

EFTASPAIN(EFTA-Spain FTA): *:*[1980-1985 (F,S)]

SADC (Southern African Development Community): *[1980-1988 and 1990-2009 (F)]*

SPARTECA (South Pacific Regional Trade and Economic Cooperation Agreement): *[1981-2009 (F)]*

CER (Australia-New Zealand FTA): *[1983-2009 (F)]*

USISR(US-Israel FTA): *[1985-2009 (F,S)]*

USCAN(US-Canada FTA): *[1989-2009 (F,S)]*

NAFTA (North American Free Trade Agreement): *[1994-2009 (F,S)]*

EC-Andorra FTA: *[1991-2009 (F)]*

EFTA-CEEC FTA: *[1992-2006 (F)]*

[37] http://www.wto.org/english/thewto_e/gattmem_e.htm
[38] http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx

EU-CEEC FTA: *[1992-2006 (F)]*

ASEAN (Association of South East Asian Nations FTA): *[1992-2009 (F,S)]*

CEFTA (Central European FTA): *[1993-2009 (F)]*

CIS (Commonwealth of Independent States): *[1995-2009 (F)]*

EAEC (Eurasean Economic Community): *[1997-2009 (F)]*

CEZ (Common Economic Zone): *[2004-2009 (F)]*

SAFTA (South Asian Free Trade Arrangement): *[2006-2009 (F)]*

WAEMU (West African Economic and Monetary Union): *[1996-2009 (F)]*

PAFTA (Pan Arab FTA): *[1998-2009 (F)]*

SACU (Sub Saharan South African Customs Union): *[2000-2009 (F)]*

EAC (East African Community): *[2000-2009 (F)]*

COMESA (Common Market for Eastern and Southern Africa): *[1995-2009 (F)]*

CAN (Andean Community FTA): *[1988-2009 (F,S)]*

MERCOSUR (Southern Common Market): *[1991-2009 (F,S)]*

DOMCAUSA (Dominican Republic - Central America - US FTA): *[2006-2009 (F)]*

TRANSPAC (Trans-Pacific Strategic Economic Partnership FTA): *[2006-2009 (F)]*

EFTASACU (EFTA-SACU FTA): *[2008-2009 (F)]*

ECSYR (EC-Syria FTA): *[1977-2009 (F)]*

ECTUR (EC-Turkey FTA): *[1996-2009 (F,S)]*

ECPAL (EC-Palestinian Authority FTA): *[1997-2009 (F)]*

ECFAR (EC-Faroe Islands FTA): *[1997-2009 (F)]*

ECTUN (EC-Tunisia FTA): *[1998-2009 (F)]*

ECMOR (EC-Morocco FTA): *[2000-2009 (F)]*

ECISR (EC-Israel FTA): *[2000-2009 (F,S)]*

ECSAFR (EC-South Afica FTA): *[2000-2009 (F)]*

EFTATUR (EFTA-Turkey FTA): *[1992-2009 (F,S)]*

EFTAISR (EFTA-Israel FTA): *[1993-2009 (F,S)]*

EFTAPAL (EFTA-Palestinian Authority FTA): *[1999-2009 (F)]*

EFTAMOR (EFTA-Morocco FTA): *[2000-2009 (F)]*

C Robustness checks for instrumented prices

This appendix shows that our results on the evolution of the substitutability parameter in the estimation with instrumented prices are robust to increasing the number of periods s for which the growth rate in the relative price of the exported composite good is predicted on the evolution of relative domestic prices.

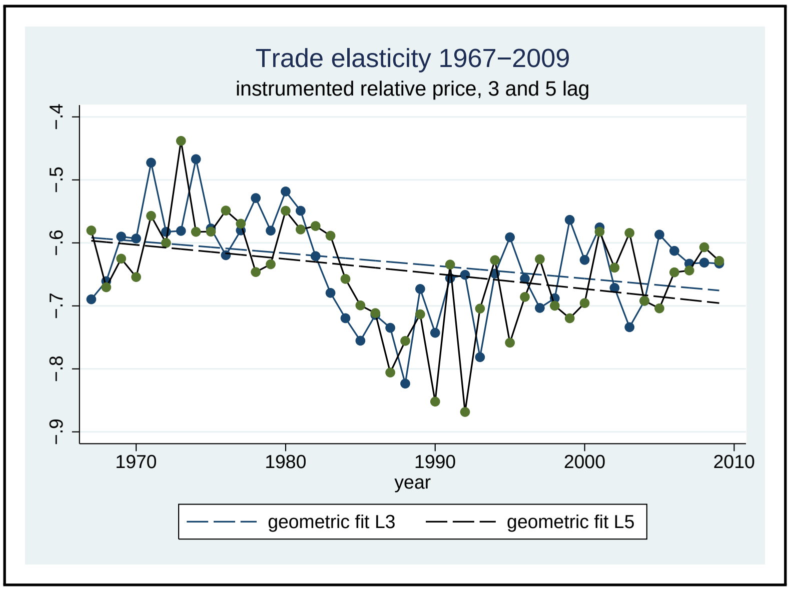


Figure 14: Estimated $(1-\widetilde{\sigma})$, instrumented relative price of composite good

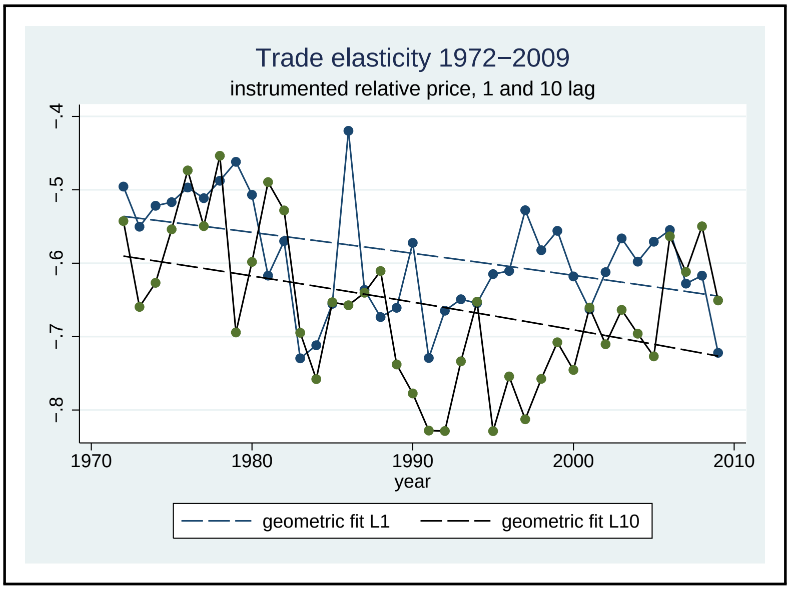


Figure 15: Estimated $(1-\widetilde{\sigma})$, instrumented relative price of composite good

Fig.(14) and Fig.(15) show that as we increase the number of lags, the evolution of the parameter becomes steeper and the level of the parameter relatively to the non-instrumented specification becomes more stable across the years. Thus, for $s = 5$, the level of the parameter increases by 20% on average relatively to the estimate obtained with observed relative prices while for $s = 10$ the level increases by 22%.

In terms of the slope, we find that for example in 1972-2009, the substitution elasticity increases by 23% when the instrument is constructed with 10 lags while it increases by 20% in the specification with just one lag. Similarly, for 1967-2009, the slope increases as we increase the number of lags from 3 to 5.