Interpreting the distance coefficient

## What does the trade elasticity actually measure?

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

In (Eaton et Kortum 2002) the heterogeneity dimension captured by the trade elasticity $\zeta $ 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 $\zeta $ is intrasectoral. Consumers value firm-specific varieties of sectoral goods which they acquire in monopolistic competition markets.

In (Anderson et 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 $\zeta $ is the lower tier Armington elasticity of substitution. It measures the degree of substitutability of goods of different national origin.[[1]](#footnote-1)

## 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 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 et Kortum 2002; Simonovska et Waugh 2014; Costinot, Donaldson, et Komunjer 2012; Caliendo et Parro 2012), but only (Levchenko et 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 et 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.

## A method to measure trade elasticity in the Armington framework

Measuring Armington elasticity is an perennial issue in trade litterature. Feenstra et al. (2014) discusses the difficulties and a state-of-the art method to measure both higher-tier (or "macro") Armington elasticities and lower-tier (or "micro") ones (see also Feenstra (1994), and refinements in Broda and Weinstein (2006) and Imbs and Méjean (2013)). We depart from this litterature and suggest a cruder estimation method for numerous reasons. First, as we want to measure the lower-tier Armington elasticity from 1963 for all the countries involved in world trade, we operate in a data-poor environement. We cannot use dissagregated domestic prices nor production. We believe that the biases entailed by this lack of information do not make the exercice worthless because we are interested in the effect of the evolution of the parameter rather than in its exact value. Second, Feenstra’s method and its developments relies on the assumption the elasticicy parameter remains constant through time, whereas our research question implies that the parameter can vary from year to year. Third, 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. While we admit that in most tracktable theoretical settings these would be the same. In the absrtact, is not immediate that they should be.

The UN COMTRADE bilateral trade database covers the majority of countries over 1962-2009. 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 should be able to observe some statistical regularities. The basic intuition of the method we use starts from the well-known result that assuming CES utility function in the one-good Armington framework we have:

\[{{X}\_{ij}}={{\left( \frac{{{P}\_{ij}}}{{{P}\_{j}}} \right)}^{-(\sigma -1)}}{{Y}\_{j}}\]

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* and \[{{P}\_{j}}\] is the price index in the destination and \[{{Y}\_{j}}\] total import demand in the destination.[[2]](#footnote-2) The exponent $\left( \sigma -1 \right)$ captures substitutability of country-composite goods across frameworks. It is also the aggregate trade elasticity $\zeta $ 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 et Méjean 2013) 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 country $i$ to destination country $j$ as the sum of imports from each sector $k$ where a sector corresponds to a SITC 4-digit category: ${{X}\_{ij}}=\sum\limits\_{k}{{{X}\_{k,ij}}}$.[[3]](#footnote-3) Given CES utility at the intersectoral level, sectoral demand in country in sector for imported goods is given by:



\[\begin{array}{\*{35}{l}}

{{Y}\_{k,j}} & = & {{\left( \frac{{{P}\_{k,j}}}{{{\beta }\_{k}}{{P}\_{j}}} \right)}^{1-\sigma }}{{Y}\_{j}} \\

\end{array}\]

Where ${{P}\_{k,j}}$ and ${{P}\_{j}}$ are price indexes, ${{\beta }\_{k}}>0$ is a sector-specific preference parameter, ${{Y}\_{j}}$ is total demand for imported goods, $\sigma >1$ is the elasticity of substitution between sectors.

Assume each country exports a specific national variety. Preferences within each sector $k$ between national varieties are assumed well represented by a CES utility function with the same $\sigma $ parameter as the intersectoral CES utility function. Intrasectoral demand for varieties exported by $i$ in $j$ in sector $k$ is:

\[\begin{array}{\*{35}{l}}

{{X}\_{k,ij}} & = & {{\left( \frac{{{p}\_{k,ij}}}{{{\gamma }\_{i}}{{P}\_{k,j}}} \right)}^{1-\sigma }}{{Y}\_{k,j}} \\

\end{array}\]

Where ${{\gamma }\_{i}}>0$ is a origin-country-specific preference parameter and ${{P}\_{k,j}}$ is the CES price index:

\[\begin{array}{\*{35}{l}}

{{P}\_{k,j}} & = & {{\left[ \sum\limits\_{i\ne j}{{{\left( \frac{{{p}\_{k,ij}}}{{{\gamma }\_{i}}} \right)}^{1-\sigma }}} \right]}^{1/(1-\sigma )}} \\

\end{array}\]

Defining $\frac{{{Y}\_{k,j}}}{{{Y}\_{j}}}={{\omega }\_{k,j}}$, we get:

\[\begin{array}{\*{35}{l}}

\frac{{{X}\_{k,ij}}}{{{Y}\_{j}}} & = & {{\omega }\_{k,j}}{{\left( \frac{{{p}\_{k,ij}}}{{{\gamma }\_{i}}{{P}\_{k,j}}} \right)}^{1-\sigma }} \\

\end{array}\]

Summing over all SITC 4-digit sectors:

\[\begin{array}{\*{35}{l}}

\sum\limits\_{k=1}^{K}{\frac{{{X}\_{k,ij}}}{{{Y}\_{j}}}}=\frac{{{X}\_{ij}}}{{{Y}\_{j}}} & = & \gamma \_{i}^{\sigma -1}\sum\limits\_{k=1}^{K}{{{\omega }\_{k,j}}}{{\left[ \frac{{{p}\_{k,ij}}}{{{P}\_{k,j}}} \right]}^{1-\sigma }} \\

\end{array}\]

The market share equation for aggregate bilateral trade as a function of the weighted average of sectoral relative prices of in is:



Changing notations:

\[\begin{array}{\*{35}{l}}

\frac{{{X}\_{ij}}}{{{Y}\_{j}}} & = & {{\kappa }\_{i}}\sum\limits\_{k=1}^{K}{{{\omega }\_{k,j}}}\frac{p\_{k,ij}^{1-\sigma }}{\sum\limits\_{l\ne j}{{{\kappa }\_{l}}}p\_{k,lj}^{1-\sigma }} \\

\end{array}\]

As this is a market share, it is reasonable to assume that the errors are multiplicative. We have to fit the following model on the data:

\[\begin{array}{\*{35}{l}}

\frac{{{X}\_{ij}}}{{{Y}\_{j}}} & = & {{\kappa }\_{i}}\sum\limits\_{k=1}^{K}{{{\omega }\_{k,j}}}\frac{p\_{k,ij}^{1-\sigma }}{\sum\limits\_{l\ne j}{{{\kappa }\_{l}}}p\_{k,lj}^{1-\sigma }} \\

\end{array}.{{e}^{{{\varepsilon }\_{i,j}}}}\]

Notice that we only have non-zero observations, as ${{p}\_{k,ij}}$ is only observed when there is a trade flow. We take logs to transform the errors into additive ones and estimate the following equation for each year with a non-linear least square procedure in STATA:

\[\begin{array}{\*{35}{l}}

\ln \left( \frac{{{X}\_{ij}}}{{{Y}\_{j}}} \right) & = & \ln \left( {{\kappa }\_{i}} \right)+\ln \left( \sum\limits\_{k=1}^{K}{\frac{{{Y}\_{k,j}}}{{{Y}\_{j}}}.}\frac{p\_{k,ij}^{1-\sigma }}{\sum\limits\_{l\ne j}{{{\kappa }\_{l}}}p\_{k,lj}^{1-\sigma }} \right) \\

\end{array}+{{\varepsilon }\_{i,j}}\]

This yields annual estimates of ${{\kappa }\_{i}}$ and $\sigma $[[4]](#footnote-4).

# Evolution of the Armington trade elasticity in 1963-2009

## 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 13-15%, and about 18% 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.

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 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. If no price is available at the 4-digit level, we move at the 3-digit level, etc. 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).

## 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*.[[5]](#footnote-5) 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 equation ) 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. [[6]](#footnote-6)

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

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 $\tilde{\sigma }$ 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**\*\*\*** | -0.2446**\*\*\*** |  | -0.0427**\*\*\*** | -0.2573**\*\*\*** | |
|  | (0.0001) | (0.0134) |  | (0.0001) | (0.013) | |
| year | -0.0029**\*\*\*** | -0.0020**\*\*\*** |  | -0.0033**\*\*\*** | -0.0024**\*\*\*** | |
|  | (0.0000) | (0.0001) |  | (0.0000) | (0.000) | |
| *ms*\**year* |  | 0.0001**\*\*\*** |  |  | 0.0001**\*\*\*** | |
|  |  | (0.0000) |  |  | (0.000) | |
| constant | 5.3474**\*\*\*** | 3.5852**\*\*\*** |  | 6.0976**\*\*\*** | 4.2515**\*\*\*** | |
|  | (0.0335) | (0.1372) |  | (0.0366) | (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 $\left( 1-\sigma \right)$ obtained when equation is estimated in annual cross-sections of the COMTRADE dataset. 2008 is obviously an outlier. However, even excluding 2008, the absolute value of trade elasticity has increased by 38% from 1962 to 2013. This corresponds to an annual increase of .6% per year.[[7]](#footnote-7) According the preceding discussion, this is a lower bound on the increase in the underlying substituability parameter.

Figure 7: Estimated (1−\[\tilde{\sigma }\] )



## Robustness checks

### 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 et 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−), BACI database

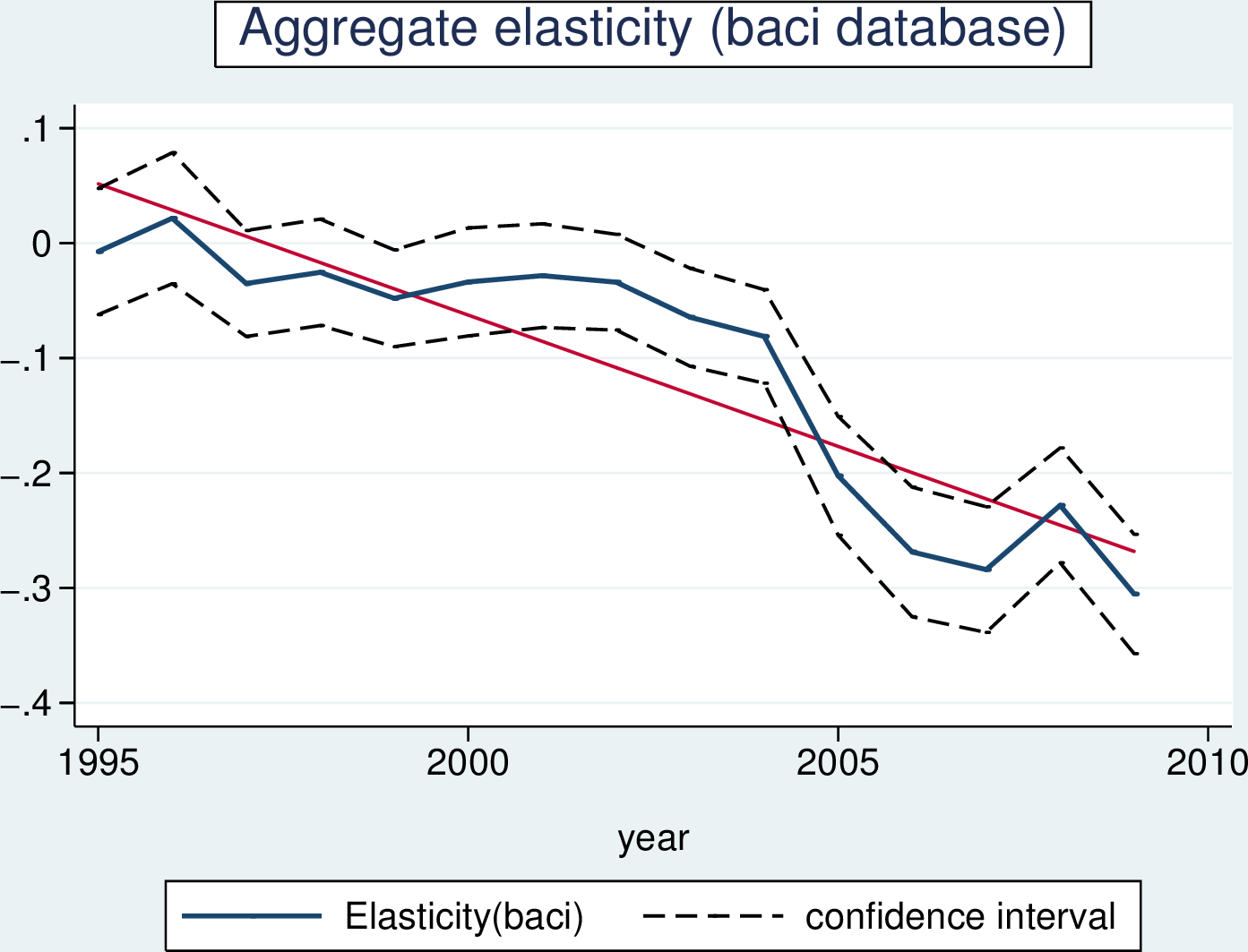


Figure 8 shows that our results hold: the elasticity parameter is found to increase in absolute value from 1995-2009. This can be compared with the equivalent period in our original dataset: the increase in the elasticity is much steeper on the BACI dataset. This finding supports the idea that our benchmark 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.

### Price instrumenting and the non-linear estimation of σ

#### Instrumenting procedure

In equation , it is possible that the estimate of $\sigma $ is downward-biased because of unobserved demand shifters that may be correlated with prices (Hummels et Schaur 2013). Specifically, whenever the supply curve has a finite slope, unobserved demand shocks will result in a simultaneous increase in the price and in the expenditure on sector *k* products. Further, even if the supply curve is perfectly elastic, measurement error may lead to a downward bias in the estimate of $\sigma $. This downward bias occurs whenever there is a systematic link between unobserved quality and observed unit values whereby higher observed unit values in sector *k* are associated with higher underlying quality and, consequently, higher expenditure on sector *k* products.[[8]](#footnote-8)

To check whether unobserved demand shifters correlated with prices are a source of concern in the estimation, we need an instrument that adequately captures exporter-specific shocks to sector *k* prices that are not demand-driven, such as exogenous shocks to the cost of inputs. We use adjusted past unit values as an instrument for current unit value. The adjustment is based on exporter-specific price shocks as measured by domestic price evolutions. Unfortunately, we do not have information on producer price indices at the SITC 4-digit level. Even at the aggregate level, PPI information is not available for most countries and years in our sample. We therefore settle for the price level of GDP and the price level of investment.

The price level of GDP will be a valid instrument if demand shocks in the importing country are the main source of price endogeneity, and the small country assumption holds. In this case, demand shocks in any market for products of a particular exporter will have no incidence on world prices of these products and, consequently, the price level of GDP in the exporting country will be independent of idiosyncratic demand shocks in the importing country. If however the small country assumption does not hold for certain importers, the price level of GDP could itself be a function of demand for exporter-specific products. Therefore, as a robustness check, we use the price level of investment as an alternative proxy for cost shocks to production because the price of investment is more likely to be determined by global demand for industrial goods and to be exogenous to unobserved demand shocks in any given market.

Information on the price level of GDP and the price level of investment is reported in the Penn World Table 9.0 for 182 countries in 1950-2015 in current US dollars (R. C. Feenstra, Inklaar, et Timmer 2015). For each variable, the price level is normalized to 1 in the USA in 2005. Information is provided for 113 countries in each year, and for the remaining countries between 10 and 51 years are covered. Thus, our sample of countries does not coincide perfectly with the countries in the Penn World Table. However, as it is the smallest exporters that drop out, the sample adjustment in terms of world trade coverage is minor.

The instrumenting procedure consists in exploiting past information on unit values and on changes in the cost of domestic inputs to replace observed unit values ${{p}\_{k,ij,t}}$ with predicted unit values ${{\hat{p}}\_{k,ij,t}}$.

We assume a stylized cost function in which the sector-specific cost is determined by an economy-wide cost measure ${{C}\_{it}}$ together with sector-producer specific characteristics summarized by the index ${{z}\_{k,i,t}}$ (> 0) and bilateral characteristics of trade costs ${{z}\_{ij,t}}$ (> 0). Denoting the time lag by $l$, we assume:

\[{{c}\_{k,ij,t}}={{C}\_{i,t}}.{{z}\_{k,i,t}}.{{z}\_{ij,t}}\]

We assume further that the cost component is well captured by the GDP or investment price level: ${{C}\_{i,t}}=P\_{i,t}^{v}$ where $v=\left\{ gdp,i \right\}$. Prices are a more or less sensitive to costs, depending on the competition conditions in each sector. The elasticity of prices to costs is \[{{\alpha }\_{k,t}}\] (>0). It is a measure of the path-through that depends on specific sector conditions. Prices are also dependent on consumer-sector-producer specific characteristics ${{z}\_{k,ij,t}}$ (> 0) that capture, for example, shocks to demand in $j$ for products from $i$ in sector $k$. As a result:

\[{{p}\_{k,ij,t}}={{\left( P\_{i,t}^{\nu }.{{z}\_{k,i,t}}.{{z}\_{ij,t}} \right)}^{{{\alpha }\_{k,t}}}}.{{z}\_{k,ij,t}}\]

Then, for each time-lag $l$, and assuming sector-specific price-setting conditions (i.e. \[{{\alpha }\_{k,t}}\]) constant between $t-l$ and $t$, we have:

\[\begin{align}

& {{p}\_{k,ij,t}}={{p}\_{k,ij,t-l}}{{\left( \frac{P\_{i,t-l}^{v}}{P\_{i,t-l}^{v}}.\frac{{{z}\_{k,i,t}}}{{{z}\_{k,i,t-l}}}.\frac{{{z}\_{ij,t}}}{{{z}\_{ij,t-l}}} \right)}^{{{\alpha }\_{k,t}}}}.\frac{{{z}\_{k,ij,t}}}{{{z}\_{k,ij,t}}} \\

& \Leftrightarrow \\

& \ln \left( {{p}\_{k,ij,t}} \right)=\ln \left( {{p}\_{k,ij,t-l}} \right)+{{\alpha }\_{k,t}}.\ln \left( \frac{P\_{i,t}^{v}}{P\_{i,t-l}^{v}} \right)+{{\alpha }\_{k,t}}.\ln \left( \frac{{{z}\_{k,i,t}}}{{{z}\_{k,i,t-l}}} \right)+{{\alpha }\_{k,t}}.\ln \left( \frac{{{z}\_{ij,t}}}{{{z}\_{ij,t-l}}} \right)+\ln \left( \frac{{{z}\_{k,ij,t}}}{{{z}\_{k,ij,t}}} \right) \\

\end{align}\]

Predicted unit values are obtained by regressing observed unit values on the product of lagged unit values and the change in the price level of domestic output (investment), assuming that the sum of the three other members of the sum follows a normal law. Our baseline equation for the first stage of the estimation for a given lag is:

\[\ln {{p}\_{k,ij,t}}-\ln {{p}\_{k,ij,t-l}}={{\alpha }\_{k,t}}\ln \left( \frac{P\_{i,t}^{v}}{P\_{i,t-l}^{v}} \right)+{{\varepsilon }\_{k,ij,t,l}}\]

We estimate this equation separately for each year of the sample. Hence, we allow the estimated coefficients to be year-specific. The identification assumption is that current shocks specific to the sector and to the pair (${{\varepsilon }\_{k,ij,t-l}}$) that may simultaneously drive up price and expenditure in year *t* are independent of lagged prices and changes in producer cost component.

The choice of the lag is associated with a trade-off. The shorter the lag, the more data we can use. However, if demand shocks in the destination are persistent, the covariance between the lagged price and the demand shock in the error term may remain positive if the first lag is used. Hence, we use a three-year lag in the baseline specifications.

For almost all years, and all specifications (there are a few exceptions at the beginning of the period) the path-through estimated in equation is statistically significant and positive. The R² is very small. That suggests that economy-wide domestic conditions explain a very small share of price and sector-specific price variations. Considering the dimension of the data, this is perhaps not surprising. Still, this may worry us as to the weakness of our instrument. Recall however that the instrument is actually the adjusted past unit value. Past unit values have a very strong correlation with present unit values. The correlation between adjusted past unit values and current unit values is slightly higher.

As a falsification test, we verify that variations of the unit price are not explained by past or future variations of domestic costs. We thus estimate the following equation:

\[\ln {{p}\_{k,ij,t}}-\ln {{p}\_{k,ij,t-l}}={{\alpha }\_{k,t}}\ln \left( \frac{P\_{i,t}^{v}}{P\_{i,t-l}^{v}} \right)+{{\beta }\_{k,t}}\ln \left( \frac{P\_{i,t+2}^{v}}{P\_{i,t+1}^{v}} \right)+\gamma \ln \left( \frac{P\_{i,t-3}^{v}}{P\_{i,t-4}^{v}} \right)+{{\varepsilon }\_{k,ij,t,l}}\]

### Instrumenting: motivation and results (old)

The results just presented are subject to caution if supply schedules are not horizontal.[[9]](#footnote-9) 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 (R. Feenstra 1994), the attenuation bias also impacts the evolution of the parameter.

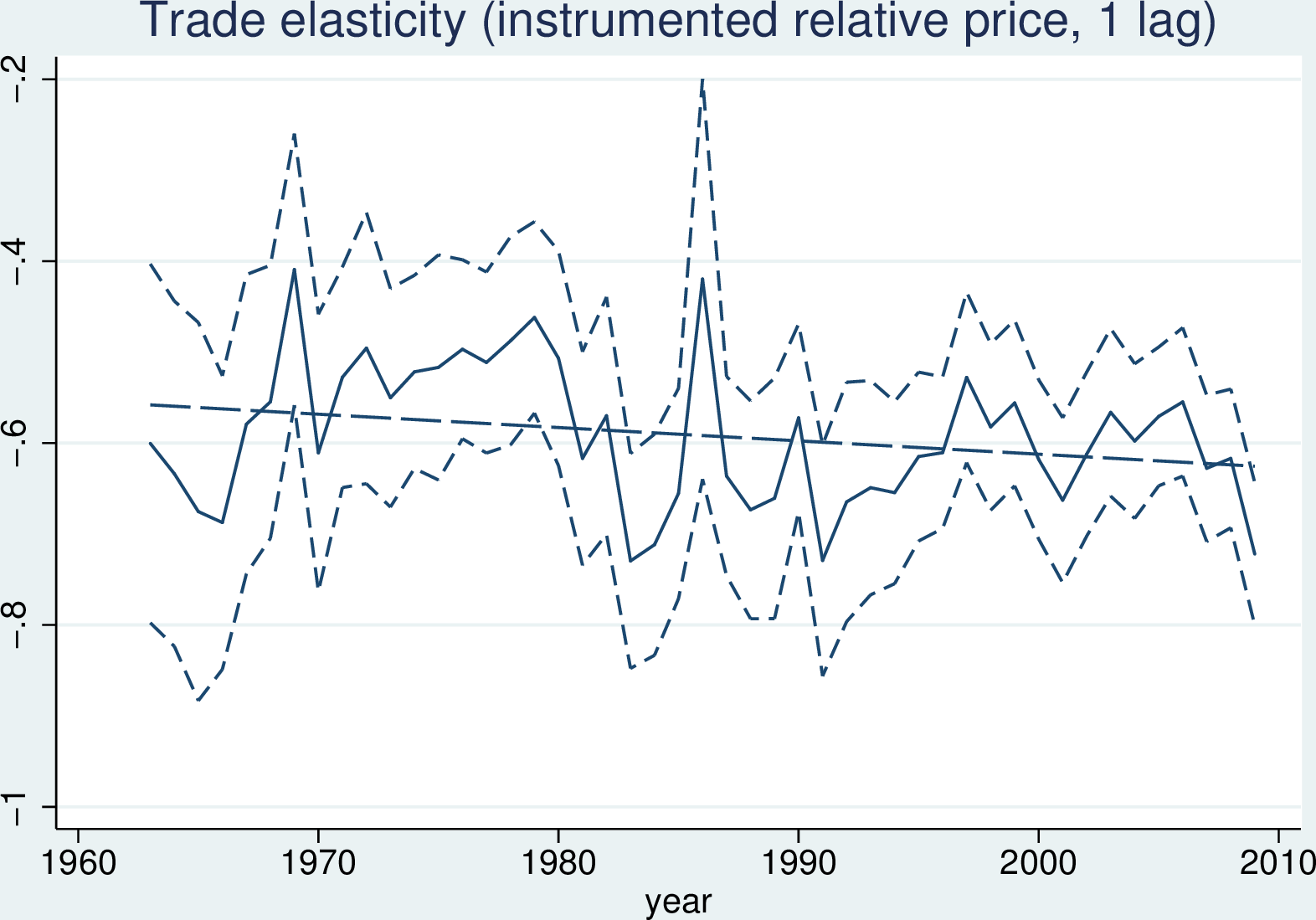
As explained in 2.3, our approach is different from the canonical one as we are keen on preserving the time dimension that is central to our analysis. We need an instrument that adequately captures exporter-specific shocks to the price of the composite good which are not demand-driven, such as exogenous 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 in 1950-2009.[[10]](#footnote-10)

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 equation 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 equation 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 Figure 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.

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



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 **Erreur ! Source du renvoi introuvable.**). 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.

## 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.[[11]](#footnote-11) 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?

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.[[12]](#footnote-12)

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 et Méjean 2013), 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 et Khandelwal 2013) as well as with increased vertical specialization of countries within sectors documented by (Fontagné, Gaulier, et Zignago 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 1962-2009, 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 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 an *ad hoc* 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, Mary, et Amit Khandelwal. 2013. « Import competition and quality upgrading ». *Review of Economics and Statistics* 95 (2): 476‑90.

Anderson, James E., et Eric van Wincoop. 2003. « Gravity with gravitas: A Solution to the Border Puzzle ». *American Economic Review* 93 (1): 170‑92.

Anderson, James E., et Eric van Wincoop. 2004. « Trade Costs ». *Journal of Economic Literature* 42 (3): 691‑751.

Arkolakis, Costas, Arnaud Costinot, et Andrés Rodriguez-Clare. 2012. « New Trade Models, Same Old Gains? » *American Economic Review* 102 (1): 94‑130.

Baier, Scott L., et Jeffrey H. Bergstrand. 2007. « Do Free Trade Agreements actually increase members’ international trade? » *Journal of International Economics* 71 (1): 72‑95.

Baier, Scott L., Jeffrey H. Bergstrand, et Michael Feng. 2014. « Economic Integration Agreements and the Margins of International Trade ». *Journal of International Economics* In Press.

Berthelon, M., et C. Freund. 2008. « On the conservation of distance in international trade ». *Journal of International Economics* 75 (2): 310‑20.

Bosquet, Clément, et Hervé Boulhol. 2009. « Gravity, log of gravity and the distance puzzle ». *GREQAM Working Paper* 2009 (12).

———. 2010. « Scale dependence of the Negative Binomial Pseudo-Maximum Likelihood Estimator ». *CES Working Papers* 2010 (92).

Broda, Christian, et Davis E. Weinstein. 2006. « Globalization and the gains from variety ». *Quarterly Journal of Economics* 121 (2): 541‑85.

Brun, Jean-François, Céline Carrère, Patrick Guillaumont, et Jaime de Melo. 2005. « Has Distance Died? Evidence from a Panel Gravity Model ». *World Bank Economic Review* 19: 99‑120.

Buch, Claudia M., Jörn Kleinert, et Farid Toubal. 2004. « The distance puzzle: on the interpretation of the distance coefficient in gravity equations ». *Economics Letters* 83: 293‑98.

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

Caliendo, Lorenzo, et Fernando Parro. 2012. « Estimates of the Trade and Welfare Effects of NAFTA ». *NBER Working Paper Series*, no 18508.

Chaney, Thomas. 2013. « The Gravity Equation in International Trade: an Explanation ». *NBER Working Paper Series* 19285.

Combes, Pierre-Philippe, Thierry Mayer, et Jacques-François Thisse. 2006. *Economie Géographique*. Paris: Economica.

Costinot, Arnaud, Dave Donaldson, et Ivana Komunjer. 2012. « What Goods Do Countries Trade? A Quantitative Exploration of Ricardo’s Ideas ». *Review of Economics and Statistics (forthcoming)*.

Daudin, Guillaume. 2003. « La logistique de la mondialisation ». *Revue de l’OFCE*, no 87: 411‑35.

———. 2005. « Les transactions de la mondialisation ». *Revue de l’OFCE*, no 92: 223‑62.

Daudin, Guillaume, Christine Rifflart, et Danielle Schweisguth. 2011. « Who Produces for whom in the world economy ». *Canadian Journal of Economics* 44 (4): 1403‑37.

Disdier, A. C, et K. Head. 2008. « The puzzling persistence of the distance effect on bilateral trade ». *The Review of Economics and Statistics* 90 (1): 37‑48.

Duranton, Gilles, et Michael Storper. 2008. « Rising trade costs. Agglomeration and trade with endogenous transaction costs. » *Canadian Journal of Economics* 41 (1): 292‑319.

Eaton, Jonathan, et Samuel Kortum. 2002. « Technology, geography, and trade ». *Econometrica* 70 (5): 1741‑79.

Egger, Peter, Mario Larch, Kevin E. Staub, et Rainer Winkelmann. 2011. « The Trade Effects of Endogenous Preferential Trade Agreements ». *American Economic Journal: Economic Policy* 3 (3): 113‑43.

Erkel-Rousse, Helene, et Daniel Mirza. 2002. « Import price elasticities: reconsidering the evidence ». *The Canadian Journal of Economics* 35 (2).

Fally, Thibault. 2012. « Structural Gravity and Fixed Effects ». *mimeo*.

Feenstra, Robert. 1994. « New Product Varieties and the Measurement of International Prices ». *The American Economic Review* 84 (1): 157‑77.

Feenstra, Robert C., Robert Inklaar, et Marcel Timmer. 2015. « The next generation of the Penn World Table ». *American Economic Review* 105 (10): 3150‑82.

Felbermayr, Gabriel J., et Wilhelm Kohler. 2006. « Exploring the intensive and extensive margins of trade ». *Review of World Economics* 142 (4): 642‑74.

Fontagné, Lionel, Guillaume Gaulier, et Soledad Zignago. 2008. « North-South competition in quality ». *Economic Policy* 23 (53): 51‑91.

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

Gaulier, Guillaume, et Soledad Zignago. 2010. « BACI: International Trade Database at the Product Level ». *CEPII working papers*, no 23.

Head, Keith, et Thierry Mayer. 2013. « What separates us? Sources of resistance to globalization ». *Canadian Journal of Economics* 46 (4): 1196‑1231.

———. 2014. « Handbook of international economics ». In , édité par G. Gopinath, E. Helpman, et K. Rogoff. Vol. 4. Elsevier.

Helpman, Elhanan, Marc Melitz, et Yona Rubinstein. 2008. « Estimating trade flows: trading partners and trading volumes ». *Quarterly Journal of Economics* 123 (2): 441‑87.

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

Hummels, David. 2007. « Transportation costs and international trade in the second era of globalization ». *Journal of Economic Perspectives* 21 (3): 131‑54.

Hummels, David, et Georg Schaur. 2013. « Time as a Trade Barrier ». *American Economic Review* 103: 1‑27.

Imbs, Jean, et Isabelle Méjean. 2013. « Elasticity Optimism ».

Johnson, Robert C., et Guillermo Noguera. 2012. « Accounting for Intermediates: Production Sharing and Trade in Value Added ». *Journal of International Economics* 86: 224‑36.

Krautheim, Sebastian. 2012. « Heterogeneous firms, exporter networks, and the effect of distance on international trade ». *Journal of International Economics* 87: 27‑35.

Larch, Mario, Pehr-Johan Norbäck, et Stephen Sirries. 2013. « Heterogenous Firms, Globalization, and the Distance Puzzle ». *IFN Working Paper*, no 957.

Levchenko, Andrei A., et Jing Zhang. 2011. « The Evolution of Comparative Advantage: Measurement and Welfare Implications ». *NBER Working Paper Series* 16806.

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

Magee, Christopher S.P, et Stephen P. Magee. 2008. « The United States is a Small Country in World Trade ». *Review of International Economics* 16 (5): 990‑1004.

Mayer, Thierry, et Soledad Zignago. 2011. « Notes on CEPII’s distances measures: the GeoDist database ». *MPRA Working Paper Series*, no 36347.

Reinert, Kenneth A., et Clinton R. Shiells. 1991. « Trade Substitution Elasticities for Analysis of a North American Free Trade Area ». *Staff Research Study, USITC*, no 19.

Saito, Mika. 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., et Silvana Tenreyro. 2006. « The Log of Gravity ». *The Review of Economics and Statistics* 88 (4): 641‑58.

———. 2011. « Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator ». *Economics Letters* 112: 220‑22.

Sato, Kazuo. 1967. « A two-level constant-elasticity-of-substitution production function ». *Review of Economic Studies* 34: 201‑18.

Schott, Peter. 2004. « Across Product vs. Within-Product Specialization in International Trade ». *The Quarterly Journal of Economics* 119 (2): 647‑78.

Simonovska, Ina, et Michael Waugh. 2014. « The Elasticity of Trade: Estimates and Evidence ». *Journal of International Economics* 92 (1).

Welsch, Heinz. 2006. « Armington elasticities and induced intra-industry specialization: The case of France, 1970-1997 ». *Economic Modelling* 23: 556‑67.

Yi, Kei-Mu. 2010. « Can Multistage Production Explain the Home-Bias in Trade? » *American Economic Review* 100 (1): 364‑93.

1. 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 et Shiells 1991; Saito 2004). [↑](#footnote-ref-1)
2. 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 cif price variations. [↑](#footnote-ref-2)
3. When several quantity units are observed, the sector is defined at the product\*quantity-unit level. [↑](#footnote-ref-3)
4. Because of the way the procedure *nl* works in STATA, instead of minimizing $\sum\limits\_{i,j}{{{\left( {{\varepsilon }\_{ij}} \right)}^{2}}}$, it minimizes $\sum\limits\_{i,j}{{{n}\_{ij}}{{\left( {{\varepsilon }\_{ij}} \right)}^{2}}}$ where ${{n}\_{ij}}$ is the number of 4-digit sectors in which imports from $i$ to $j$ are recorded. As a result, we have to weight the regression by $1/{{n}\_{ij}}\ $ to make its results comparable to those in part one. [↑](#footnote-ref-4)
5. 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. [↑](#footnote-ref-5)
6. 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. [↑](#footnote-ref-6)
7. The coefficient of the geometric fit is significant at 1% level, the 95% confidence interval is between 0.4% and 0.9% (not taking into account though the uncertainty around yearly estimates). [↑](#footnote-ref-7)
8. Another potential problem is linked to using per kg instead of per item prices. As discussed in (Hummels et Schaur 2013), unit values reflect product bulkiness, and bulkier products are more likely to be shipped via cheaper means of transportation. If products differ in bulkiness within the sector, unit value differences will overstate price differences and lead to a downward bias in the estimated elasticity. [↑](#footnote-ref-8)
9. (Broda et Weinstein 2006) find that supply elasticities are finite at the 4-digit level. On the other hand, (Magee et Magee 2008) find that the small country assumption may hold in the data in which case there would be no attenuation bias. [↑](#footnote-ref-9)
10. See (Heston, Summers, et Aten 2011). Our sample of countries does not coincide perfectly with the countries in the Penn World Tables. However, as it is the smallest exporters that drop out, the sample adjustment in terms of world trade coverage is minor. [↑](#footnote-ref-10)
11. The elasticity decreases by 5% when the evolution of $\rho $ is computed from the ratio of trends, and by 7% when it is computed as the trend of the ratios. [↑](#footnote-ref-11)
12. (Schott 2004) documents increased similarity in the set of exported goods of US trade partners while (Broda et Weinstein 2006) document the increase in the number of imported varieties since the 1970s. [↑](#footnote-ref-12)