Estimating Regional Trade Elasticities in the EU Internal Market Using Generalized Transport Costs

José L. Zofío[a,1], Jorge Díaz-Lanchas[b], Damiaan Persyn[b] and Javier Barbero[b]

[a] Department of Economics. Universidad Autónoma de Madrid, Madrid, Spain.
Erasmus Research Institute of Management, Erasmus University, Rotterdam, The Netherlands.
[b] European Commission, Joint Research Centre (JRC), Seville, Spain.

Preliminary Draft, September 6, 2019

Contents

---

[1] Corresponding author: José L. Zofío. Voice: +34 914972406; E-mail: jose.zofio@uam.es, jzofio@rsm.nl.

Abstract

This paper undertakes the simultaneous estimation of import elasticities of substitution (trade elasticities) within European Union (EU) regions differentiating between imports from regions belonging to the same country (intranational or regional trade) and regions belonging to other EU countries (international trade within the EU). From the theoretical model that allows for these two alternative trade flows depending administrative boundaries, we derive the corresponding gravity equation that is amenable to econometric estimation by way of the Poisson pseudo-maximum likelihood regression. As the UE is a single market, the usual approach relying on bilateral tariffs, cannot be used to identify the trade elasticities. To address this issue a very detailed definition and calculation of the ad valorem specification of transport costs is performed. These calculations take into account the transport engineering and logistic characteristics of road freight transportation, which allows us to obtain a reliable measure of the generalized transport costs between regions. The trade elasticities are calculated at the 2 digit level categories belonging to the agriculture, mining and natural resources, and several manufacturing sectors. In our preferred specification we can recover elasticities that are statistically significant for all sectors. Results show that the trade elasticity increases the closer are the trading partners. Specifically, the average value of the trade elasticities between regional (domestic) goods and those imported from regions within the same country is 26.9, ranging from 1.8 to as much as 134.0, while that for foreign goods sourced from regions situated in other EU countries is 17.9, ranging from 2.1 to 187.2. Consequently, national trade elasticities are in general larger in magnitude that their foreign counterparts. Our calculated trade elasticities can be adopted in a wide variety of models of international trade, or spatial economic models such as Regional Computable General Equilibrium models, improving the results obtained from simulations aimed at policy analysis.

1. Introduction

Elasticities of import substitution play a key role in modern trade theory by capturing the sensitivity of consumer's relative demand for domestic and foreign goods to changes in their relative prices (Hillberry and Hummels, 2013). Under the usual assumption of constant elasticity of substitution (CES) between goods in the utility function, the elasticity between any two varieties produced in different foreign locations, corresponds to the elasticity of import substitution (hereafter, trade elasticities); i.e., the inverse of the cross-price elasticity of demand between foreign goods (Feenstra, 2016). Once embedded in a Computable General Equilibrium (CGE) framework these elasticities shape market dynamics of the output and input markets and characterize the equilibrium of the economy. For open sectors with tradable goods, they determine the quantities demanded that are domestically consumed or sourced from abroad (import trade flows), as well as the quantities that are supplied to local or foreign markets (export trade flows). The ripple effects of shocks to the trade sector emanating from changes in transport and non-transport related costs (such as infrastructure investments or tariffs or), are particularly relevant in open economies like those belonging to free trade areas or, as in the case of the EU, single markets. For example, trade liberalization brings about relevant modifications in the structure of the output and inputs markets, and the location of economic activity both between and within countries, Gallego and Zofío (2018). In the markets for goods they tend to disrupt the status quo by altering the degree of competition through changes in the size of firms, normally reinforced with selection effects, Burstein and Melitz (2013). In the input labor market, changes require longer response times as a result of rigidities and frictions. Potential welfare gains (e.g., (un)employment rates) are highly dependable on the industry workforce demographics (skills and age), e.g., Dix-Carneiro (2014). This multiplicity of interrelated effects across the economies can only be captured within a general equilibrium setting, and CGE models become key for policy analyses and evaluations.

The central role played by the trade elasticities explains the interest in obtaining reliable estimates that can be later employed as parameters in the calibration of CGE models. Major trade-focused CGE models draw elasticities from a wide range of studies. These econometric studies follow alternative specifications (gravity equations, demand or supply equations), estimation methods (ordinary least squares, Poisson pseudo maximum likelihood,...), sample data specific to geographic locations and time (e.g., world regions, particular free trade areas,…), etc. This translates in a numerous set of results and the modeler's question is what the best elasticities for the model at hand are. Although one can always find a close match between the trade oriented CGE models and econometric results, there is an area in which there has not been much headway. Specifically there exist very few studies on trade elasticities between regions belonging to the same country, or, if data is available as in our case, between regions belonging to several countries. Indeed, most of studies surveyed in literature reviews (e.g., Francois and Martin, 2013; Hillberry and Hummels, 2013) refer to international GCE modelling, where trade takes place between countries and there is a single trade elasticity parameter capturing the relationship between either domestic and foreign goods or between goods sourced from different countries (foreign-foreign). There are, however, a handful of CGE models for single countries, where regions trade with each other. In these models the transport related costs play a leading role in the identification and quantification of trade elasticities since changes in relative import prices are not driven by shocks to tariffs or any other non-transport related cost. Hence, to estimate trade elasticities within countries ones needs to rely on intra-national (i.e., interregional) trade flows and the existing relationship with actual transportation costs (Hilberry and Hummels, 2008, Díaz-Lanchas et al., 2019).[2] Arguably, the estimation of trade elasticities is more challenging for single market CGE models than for their internationally oriented counterparts, due to the fact that the difference between domestic and imported goods is conceptually blurry, and cannot be determined in terms of shocks to standard non-trade related barriers (e.g., tariffs)—*à la* Hertel et al. (2007).

Whether the model includes countries only or incorporates regions is key to the identification and numerical determination of trade elasticities. Adopting the approach that relies on the specification of gravity equations with a microeconomic theory foundation, the value of the trade elasticities can be identified and, therefore, recovered from the estimates of the coefficients associated to trade costs (i.e., the source of price variation among varieties). Transport costs are typically expressed on a multiplicative (i.e., iceberg or ad valorem), or an additive (i.e., per unit) basis (see, e.g., Irarrazabal et al., 2015), such that the price in the destination including all transport costs either equals either the price in the origin *times* the proportion (τ) corresponding to trade costs in the multiplicative case, $p(1+\tau)$, $\tau \geq 0$, or the price in origin *plus* transport costs, $p + t$, $t \geq 0$, for the additive case (hence, independent of the price in origin).[3] In CGE models the former definition is normally adopted to prevent an additional source of non-linearity in the equilibrium equations, which eases convergence. Therefore prices at the destination are equal to mill prices in the origin times trade costs, expressed as an ad-valorem freight rate.

In models consisting only of countries do not belonging to a single-market, the trade costs of interest from a modeling perspective are tariffs, whose level is customarily changed to assess the effects of polices aiming at increasing or reducing trade liberalization. Transport related costs are of secondary importance and can be fairly regarded as control variables, whose measurement does not require extraordinary detail or complexity, and therefore may be adequately represented by average freight rates between countries, or even simple geographic distance.

---

[2] In passing we note that this by-passes one of the problems identified by Hilberry and Hummels (2016) related to the correlation between non-transport related costs such as tariffs and the error term in the gravity equation. I.e., political economy suggests that tariffs are higher when the threat of potential import substitution is higher.

[3] Although it is obvious that the iceberg or ad valorem definitions can be set to match any observed value of transport costs, the functional form starts to matter when the price changes. Imagine that a production tax or an increase in quality leads to higher prices of some tradable goods. Under the assumption of multiplicative transport costs the transport costs would increase proportionally (except perhaps for the insurance premiums); an effect that would most likely be unintended and not realistic.

This is however not the case for regional CGEs oriented towards a lower level of spatial disaggregation, normally characterized by single markets and where tariffs have been removed. These models situate between international models with many countries (e.g., Global Trade Analysis Project, GTAP) and single country models (e.g., US International Trade Commission, USAGE). Central to our study, a representative example of this type of models is the RHOMOLO model for the European Union (EU).[4] The RHOMOLO model draws from previous experiences of regional EU CGE modelling (e.g., Bröker, 2015), and is maintained by the REMO modelling team of the Joint Research Center of the European Commission. The current—third— version of the model features the most relevant and latest advances in regional modelling and trade theory (e.g., competitive and imperfectly competitive markets, alternative labor market closures allowing for rigidities and frictions in the wage curve, etc…)—see Lecca et al. (2018). It is arguably the largest and most complex model in terms of its spatial dimension by covering a total of 267 NUTS-2 regions within 28 countries, disaggregating their economies into ten NACE rev. 2 sectors.

Relaying on the theoretical features of this model and its associated trade data, the most salient contribution of this paper is that, by taking advantage of the trade data corresponding to the three geographical levels represented in the RHOMOLO model: regional, national and international, we can estimate two import elasticities of substitution. Relying on the customary Armington assumption that buyers treat varieties as differentiated on the basis of the location of origin we distinguish, on one hand, the elasticity of substitution between domestic goods and those imported from abroad, and originating in foreign regions located in different EU countries–as in the existing studies for international trade. On the other hand, we exploit the information on trade between a given region and other regions *within* the same country to estimate a second import elasticity of substitution for imports, which are supposed to be closer in the product space given that they share similar idiosyncratic characteristics and are better known to consumers.[5] On these grounds it is hypothesized that for the EU, international trade elasticities (i.e., between regions in different countries) represent lower bounds for regional trade elasticities (i.e., between regions in the same countries). Our results confirm this hypothesis, concluding that the closer the geographical reach of trade, i.e., regional versus international trade, the higher the substitutability between domestic and imported varieties.

We compare the results obtained with previous estimates of trade elasticities at both the country and regional level. Most of the elasticities based on the Armington assumption refer to the US and very few to Europe. At the country level, and for the European case there are some recent econometric estimates by Németh et al. (2011), Olekseyuk and

---

[4] https://ec.europa.eu/jrc/en/rhomolo.

[5] In the literature there is the distinction between home-foreign substitution and foreign-foreign substitution, where the former is obtained from time-series data referring to the same imported across-time, and the latter would be obtained from cross-sectional data as in the present study (see, e.g., Németh et al., 2011). Moreover, when cross-section estimates for CGE multicounty modeling where not available the "rule of two", by which foreign-foreign substation was twice the values of the home-foreign elasticity of substitution, was generally applied, Hilberry and Hummels (2016, 1128).

Schürenberg-Frosch (2016) and Aspalter (2016). The range of elasticities in each of these studies go from around 2 to 5, in the interval of 3 and 4.2, and 0.3 and 3.7, respectively. These elasticities appear to be consistent with other studies where single European countries (i.e. regional elasticities) are considered (Welsch, 2008; Imbs and Méjean, 2010 and 2015). As expected, these studies reveal that trade elasticities exhibit a great deal of heterogeneity between countries, but also depending on the level of industrial aggregation; i.e., as would be the case when moving upwards in the digit classification by successively increasing the aggregation of trade flows (e.g., from a three to a two-digit classification). In this respect, as reported by Aspalter (2016) for the European case the difference between 'micro-elasticities' and 'macro elasticities' can be significant, with the former exhibiting lower values than the latter. This difference exists, to even a greater, in the US as shown by Hummels (2001), Feenstra et al. (2014) and Imbs and Méjean (2015).

Key to the estimation of trade elasticities, whose values we take particular care when calculating, is a very detailed matrix of transport costs. Rather than assuming proxies, the definition of trade costs follows a generalized transport cost (GTC) approach, which calculates the minimum cost of shipping freight between any two locations along the least expensive route. Given the percentage of freight transportation in the EU by road (over 85%), and the impossibility of setting an EU wide intermodal freight transportation model due to the lack of reliable statistics, we focus on the road transportation mode.[6] The methodology relies on a computationally intensive process that takes into account the economic costs of transportation, and where the choice of the optimal vehicle size depends on: a) 'freight curves', balancing fixed costs such as terminal times (handling costs) and variable costs (hauling costs) (McCann, 2001); b) the urban layout of the origin and destination, and c) the type of commodity transported. Here the transport engineering and logistics approach presented by Zofío et al. (2014) is enhanced to account for the existence of non-linear shipping costs associated to economies of distance and size. The aim is to reduce to a minimum the likely correlation between transport related costs and the error term in the gravity equation, by making sure that the ad-valorem transportation costs control for all this specificities. This information is then embedded in a geographical information system (GIS) representing the digitalized transportation network across the EU. Following Persyn et al. (2018), the optimal route associated to the minimum cost is calculated by using Dijkstra's algorithm. A sample of centroids based on nighttime satellite imaging at a 1kmx1km resolution is considered for every pair of regions, in order to take into account the typically uneven distribution of economic activities within the regions. Subsequently, the set of bilateral GTCs between any two regions is aggregated through the harmonic mean, as in Head and Mayer (2003). In the last step, the average GTC is employed to calculate the ad valorem cost of transporting the observed trade between any two regions, which is the variable that is used in the estimation of the gravity equation.

---

[6] Since 1999 non-road transportation modes, mainly train and inland waterways, has stalled. Mostly due to a low containerization rate, deterioration in the quality of services of intermodal transport, and improvements in the efficiency and quality of road transport services, Janic (2007).

The gravity equation is derived from a theoretical model that follows the RHOMOLO framework, whose characteristics, beyond EU specificities, are nevertheless common to earlier EU regional CGEs, such as Bröcker (2015). Although the underlying assumptions could be easily changed, as the estimated trade elasticities will enter the set of parameters needed to calibrate the model, we rather maintain the basic framework. This should results in more reliable simulations exercises upon which regional policy analyses are based—see Lecca et al. (2018) for a general description of this model and Di Comité et al. (2018) for different policy impact assessment within a New Economic Geography framework. This means that household and firm behavior follow standard specifications. The representative consumer maximizes a three-tier utility function where the middle and lower tier is characterized by an asymmetric CES utility and substitutability is constant among varieties. As for markets, these are characterized by either constant or increasing returns, resulting in perfect and imperfect competition, and where the market structure in the latter case is assumed to be that of monopolistic competition, thereby doing away with strategic behavior among firms.

The paper is structured as follows. In the next section we present the theoretical model underlying the specification of the gravity equation. Here we derive the demand equations for domestic goods, intranational imported goods within the same country, and internationally imported goods from regions located in foreign countries. Next, in the third section, we discuss the specific econometric specification of the gravity equations and the estimation strategies. In the fourth section we discuss the data related to trade flows, generalized transport costs, and ancillary control variables. Here we introduce the methodology employed to calculate the ad valorem trade costs based on the generalized transport costs that taking into consideration the choice of optimal vehicle size depending of three factors: distance between the origin and destination location, their relative degree of urbanization and the nature of the commodity transported. In section five we present our estimates of the trade elasticities differentiating between interregional and international substitutability, as well as the level of industrial (sectoral) aggregation. Finally, we conclude with several ideas regarding the novelty of results and their relevance for regional CGE modelling.

2. The model: A gravity specification

The theoretical model from where we derive the import demand equations underlying the gravity equation for non-domestic goods is consistent with the regional CGE RHOMOLO model. In this framework household preferences are modelled through a triple nested specification of the utility function. The upper tier corresponds to

$$U = \left(Q_1, Q_2, ..., Q_c, ..., Q_C\right) \ , \tag{1}$$

which aggregates the *c* = 1,…,*C* quantities of commodities demanded by the representative consumer (normally aggregated into sectors based on their similarity for statistical convenience; e.g., agriculture, manufacturing, services,…), and whose functional form may range from the simplest Cobb-Douglass formulation, the Constant Elasticity of Substitution (CES), the quasi-linear or quadratic specifications, to more complex non-homothetic characterizations if income effects are of interest (Fieler, 2011).

For open sectors the amount consumed is a composite of horizontally differentiated varieties of the same good that may be produced locally (domestic consumption), or imported either from regions from within the same country, or from regions situated in foreign countries. With this structure in mind, the middle tier of the utility function is expressed by way of the following CES specification:

$$Q_d^c = \left[b_{Dd}^c {Q_{Dd}^c}^{\frac{\phi^c-1}{\phi^c}} + b_{Nd}^c {Q_{Nd}^c}^{\frac{\phi^c-1}{\phi^c}} + b_{Fd}^c {Q_{Fd}^c}^{\frac{\phi^c-1}{\phi^c}} \right]^{\frac{\phi^c}{\phi^c-1}}, \tag{2}$$

where Q_d^c is the total quantity of the composite good of sector *c* consumed in region *d,* which as previously stated can be domestically produced or imported from other regions. Among the latter we make a further distinction by differentiating between intranational and international trade, and, therefore, Q_d^c is the result of aggregating the composed domestic good (*D*), Q_{Dd}^c, the composed imported good from regions within the same country (*N*), Q_{Nd}^c, and the composed imported good from regions from other countries (*F*), Q_{Fd}^c. In (2), the parameters b_{Dd}^c , b_{Nd}^c and b_{Fd}^c represent preference weights specific to each source; i.e., domestic, nationally imported and internationally imported. The parameter ϕ_c > 1 represents the elasticity of substitution among these alternative sources of good *c* in region *d.* We assume that this elasticity is equal across the European Union *d* importing regions.

Finally, in the lower tier, and denoting by *n*=1,…,*N* and *f*=1,…,*F*, the exporting regions belonging to the same country of the importer and those situated abroad, respectively, the composite demands for goods having a national (*n*) or foreign (*f*) origin are represented by their corresponding CES functions:

$$Q_{Nd}^c = \left[\sum_{n=1}^{N} b_{nd}^c {q_{nd}^c}^{\frac{\sigma_N^c-1}{\sigma_N^c}} \right]^{\frac{\sigma_N^c}{\sigma_N^c-1}}, \tag{3}$$

$$Q_{Fd}^c = \left[\sum_{f=1}^{F} b_{fd}^c {q_{fd}^c}^{\frac{\sigma_F^c-1}{\sigma_F^c}} \right]^{\frac{\sigma_F^c}{\sigma_F-1}}, \tag{4}$$

where q^c_{nd} and q^c_{fd} are the quantity of sector c commodity consumed in *d* imported from regions in the same country n, and from regions in other countries f, respectively. In this level b^c_{nd} and b^c_{fd} represent the corresponding preference parameters for each one the varieties imported from the national or foreign regions, and σ_N and σ_F are the associated elasticities of substitution among varieties sourced from each group of regions. Once again, we assume that these elasticities are equal across the European Union *d* importing regions.

We now determine the aggregate demand of the national and international imported goods for the representative consumer maximizing (2) conditional on the expenditure on each type of commodity, $E^c_{di}, i = N, F$ (coming from the upper level utility function (1)), and assuming that the relevant market structure corresponds to monopolistic competition.[7] In this case the optimal sourcing of imports from different importers, *n* or *f*, according to (3) and (4), results in the following demand equations:

$$q^c_{nd} = {b^c_{nd}}^{\sigma^c_N} \frac{{p^c_{nd}}^{-\sigma^c_N}}{{P^c_{Nd}}^{1-\sigma^c_N}} E^c_{Nd}, \tag{5}$$

$$q^c_{fd} = {b^c_{fd}}^{\sigma^c_F} \frac{{p^c_{fd}}^{-\sigma^c_F}}{{P^c_{Fd}}^{1-\sigma^c_F}} E^c_{Fd}, \tag{6}$$

In these commodity specific demands, destination prices in the numerator correspond to the following specifications:

$$p^c_{nd} = \left(\frac{\sigma^c_N}{\sigma^c_N - 1}\right) p^c_n (1 + \tau^c_{nd}), \text{ where } p^c_n = c^c_n w^c_n, \text{ and} \tag{7}$$

$$p^c_{fd} = \left(\frac{\sigma^c_F}{\sigma^c_F - 1}\right) p^c_f (1 + \tau^c_{fd}), \text{ where } p^c_f = c^c_f w^c_f\ . \tag{8}$$

In these expressions $p^c_n = c^c_n w^c_n$ and $p^c_f = c^c_f w^c_f$ are mill prices in origin depending on the national or foreign region (represented by a constant marginal labor requirement c^c_i multiplied by wages in origin, w^c_i, *i* = *n*, *f*; i.e., the marginal cost);[8] τ^c_{nd} and τ^c_{fd} are *ad valorem* transport costs between the exporting and importing region; and $\sigma_N/(\sigma_N - 1)$, and

---

[7] In the RHOMOLO model sectors are assumed to be either perfectly competitive or characterized by monopolistic competition—see Lecca (2018: 11). Since the final econometric specification that allows the identification of the trade elasticities associated to the trade costs does not differ between the two, we show the general case corresponding to monopolistic competition. I.e., the equilibrium condition under perfect competition corresponds to the simplest case where marginal revenue equals price, and therefore the profit maximizing condition for the firms requires that prices equal marginal cost.

[8] Note that there is a single mill (factory gate) price in each exporting regions, and therefore it is assumed that they do not undertake price-discrimination depending on the region of destination. This is consistent with EU competition law, banning this discriminatory practices. The legal definition of price discrimination in Article 102 (TFEU) refers to the application of "*dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage*". In international trade agreements this corresponds to the most favored nation clause.

$\sigma_F/(\sigma_F - 1)$ are the mark-ups reflecting the degree of market power under monopolistic competition.

Finally, the overall price index over the imported commodities are:

$P_{Nd}^{c} = \left(\sum_{n=1}^{N} {b_{nd}^{c}}^{\sigma_N} (p_{nd}^{c})^{1-\sigma_N}\right)^{1/1-\sigma_N}$, and

$P_{Fd}^{c} = \left(\sum_{f=1}^{F} {b_{fd}^{c}}^{\sigma_F} (p_{fd}^{c})^{1-\sigma_F}\right)^{1/1-\sigma_F}$.

3. Econometric specification and estimation of trade elasticities.

The econometric specification of the gravity equations that allows identifying the elasticities of import substitution relies on the cross-sectional variation of delivered prices induced by trade costs. In our single market setting characterizing the EU, delivered prices corresponds to mill prices plus the trade margins, of which ad valorem transport costs represent the largest proportion, and excluding non-transport related costs since there are no additional trade barriers such as tariffs.[9] Given the available information, and following standard practice, this requires expressing the demand equations (5) and (6) in value terms by multiplying both sides by destination prices. Also in a monopolistic competition framework the aggregate import value can be related to the individual firm *h* exports multiplied by the number of symmetric firms *m* operating in the exporting industry; i.e., $V_{id}^{c} = p_{i}^{c} m_{i}^{c} q_{hid}^{c} = p_{i}^{c} q_{id}^{c},\ i = n, f$.[10] Then, multiplying (5) by (7) as presented in the second equality, and taking natural logs of the resulting equation, yields the following gravity equation for intranational trade:[11]

$$\begin{aligned}\ln V_{nd}^{c} = \sigma_{N}^{c} \ln b_{nd}^{c} + \ln m_{n}^{c} + (1-\sigma_{N}^{c}) \ln\left(\frac{\sigma_{N}^{c}}{\sigma_{N}^{c}-1}\right) + (1-\sigma_{N}^{c}) \ln(c_{n}^{c} w_{n}^{c}) \\ + (1-\sigma_{N}^{c}) \ln(1+\tau_{nd}^{c}) + (\sigma_{N}^{c}-1) \ln P_{Nd}^{c} + \ln E_{Nd}^{c}\end{aligned}$$

In the same vein, multiplying (6) by (8) it is possibility to obtain the gravity equation for international trade:

---

[9] The difference between (export) FOB and (import) CIF definitions of trade flows, with the latter including not only transport costs, but also insurance and taxes, becomes relevant when compiling the data for the empirical estimations. This is discussed in section 4.1 below.

[10] We assume that all firms within a given region operate with the same technology and face the same inputs costs. Consequently, for simplicity, we drop the firm specific subscript *h* in the following expressions.

[11] In the final econometric specifications of the gravity equations shown below the number of firms or varieties, along with the preference parameters, and any origin-specific determinants are eventually swept out by the fixed effects capturing export-only characteristics. Correspondingly, the importer region's price index, expenditure, and any other destination-specific determinants are also swept out by the importers' fixed effects.

$$\ln V_{fd}^{c} = \sigma_F^c \ln b_{fd}^{c} + \ln m_f^c + (1-\sigma_F^c)\ln\left(\frac{\sigma_F^c}{\sigma_F^c - 1}\right) + (1-\sigma_F^c)\ln\left(c_f^c w_f^c\right) + (1-\sigma_F^c)\ln(1+\tau_{fd}^{c}) + (\sigma_F^c - 1)\ln P_{Fd}^{c} + \ln E_{Fd}^{c}$$

3.1. Individual (intra)national and foreign (international) trade elasticities by sector.

The above specifications are amenable to individual econometric estimation by the nature of the trade flows (either intra- or international) and sector. The standard econometric strategy followed by authors such as Hummels (2001) and Hertel *et al.* (2007) exploits the fact that all variables except the bilateral preferences and transportation costs: b_{id}^{c} , $\tau_{id}^{c}, i = n, f$, are either importer or exporter specific, and therefore their effects on bilateral trade can be captured though their individual fixed coefficients. Denoting by a_d^c and a_n^c the vectors of importer and exporter (within the same country) regional fixed effects, results in the following specification:

$$\ln V_{nd}^{c} = a_d^c + a_n^c + \sigma_N^c \ln b_{nd}^{c} + (1-\sigma_N^c)\ln(1+\tau_{nd}^{c}),\ c{=}1,\ldots,C. \tag{9}$$

while the international counterpart corresponds to the following specification.

$$\ln V_{fd}^{c} = a_d^c + a_f^c + \sigma_F^c \ln b_{fd}^{c} + (1-\sigma_F^c)\ln(1+\tau_{fd}^{c}),\ c{=}1,\ldots,C. \tag{10}$$

The above expressions deal with simultaneity and mis-measurement problems by assuming that prices, quality and other commodity characteristics are the same in each import destination to which an exporter $i\ (=\ n, f)$ sells to, and can therefore be captured by the fixed effects. Here the bilateral variation in the commodity-specific preference parameters b_{id}^{c} captures other idiosyncratic characteristics that may foster trade between the importer and exporter. In particular, it is customary to include distance and adjacency (contiguity), when defining this variable: $b_{id}^{c} = b_i^c Dist._{id}^{\delta_1^c}\ b_i^c e^{\delta_2^c Adj.}$.

Rather than estimating the sector specific elasticities of trade for national and foreign goods separately, i.e. using subsamples corresponding to each type of trade flow, our estimation strategy uses the whole combined sample of trade flows (i.e., by pooling the two levels of trade). This implies that a single specification of the gravity equation can be implemented in such way that σ_N^c and σ_F^c can be recovered from the estimated parameters. This is achieved by defining an specifications that allows for both levels of trade flows, i.e., regressing all trade flows on $\tau_{id}^{c}, i = n, f$, but subsequently qualifies this overall value by introducing a variable that controls for intranational trade. This variable is simply defined as the interaction between transportation costs and a dummy capturing if the flow is intranational: i.e., $\tau_{nd}^{c} = \tau_{id}^{c} \times D_{nd}^{c}$, where $D_{nd}^{c} = Bord.country$, and its associated parameter effectively captures the additional (marginal) effect on imports if trade is intranational rather than international (i.e., the reference category). The latter becomes the model baseline, and the corresponding results are identical to those that would be obtained if the intranational and international equations (9) and (10) were run independently. Considering this estimation strategy results in the following specification:

$$\begin{aligned}\ln V_{id}^{c} = \alpha_0 &+ \alpha_d^c + \alpha_i^c + \beta_i^c \ln(1+\tau_{id}^c) + \beta_n^c \ln(1+\tau_{nd}^c) + \beta_1^c \ln Dist._{id} + \\ &+\beta_2^c\, Bord.region_{nd} + \beta_3^c\, Bord.country_{fd} + \beta_4^c\, Adj.region_{nd} + \beta_5^c\, Adj.country_{fd} \\ &i = n,f,\ \ c = 1,\dots,C.\end{aligned} \tag{11}$$

Here α_d^c and α_i^c are the import and export specific parameters capturing the corresponding fixed effects, $Bord.region_{id}$ and $Bord.country_{id}$ are the classic definitions of a pair of dummy variables capturing border effects if the trade flows take place within the same region and/or within the same country (as above), while $Adj.region_{id}$ and $Adj.country_{id}$ are dummy variables reflecting if the flows take place between adjacent regions or countries. The foreign and national elasticities of trade can be identified from the parameters associated to the bilateral variation in transportation costs: i.e, $\sigma_F^c = 1 - \beta_i^c$ and $\sigma_N^c = 1 - (\beta_i^c + \beta_n^c)$. In the results section where we present the values of the trade elasticities according to expression (11), the statistical significance of these parameters is critical. For example, if any of the two is non-significant, then we can conclude that there is no difference between the two, while if both are significant, they will differ according to the value of β_n^c (clearly, if both are non-significant, no trade elasticity can be recovered from the data). Precisely, the advantage of relying on a combined specification such as (11) with respect to the alternative individual regressions is the possibility of testing whether the gap between the national and foreign trade elasticities is statistically significant or not. As for the estimation method we follow Santos Silva and Tenreyro (2006, 2010, 2015) and Francois and Manchin (2013) in relying on the Poisson Pseudo Maximum Likelihood (PPML) method as the most suitable econometric approach. The Poisson estimator is consistent and unbiased in presence of heteroscedasticity when the data have a large number of zeros. Additionally, it yields more efficient estimators than their OLS counterparts. Additionally, this method identifies and eventually drops regressors that may cause the non-existence of the (pseudo) maximum likelihood estimates, presenting several advantages given the problems posed by the existence of numerous zeros and use of dummy variables (see also Head and Mayer, 2013).

3.2. Pooling (intra)national and foreign (international) trade data by sector.

As it is normal practice in the literature, the above regressions yield estimates of sector specific trade elasticities (either intranational or international) that are individually obtained for each sector. However, it is also possible to pool the data into a single regression relying on sector specific dummies that, on top of the dummy identifying whether the trade flows are intranational as above, captures the particularities of trade flows in commodities belonging to the same sector. Thus, we define $\tau_{id}^c = \tau_{id} \times D_{id}^c$, *c* =1,…, 19 are the actual dummies for each sector, while $\tau_{nd}^{c} = \tau_{id} \times D^{c} \times D_{nd}^{c}$ (where, again, $D_{nd}^{c} = Bord.country$) now identifies whether a sector specific trade flows takes is intranational by taking place between region within the same country. In this way, pooling all trade data into a single specification ensures that the import and export commodity-specific fixes effects are kept equal regardless of level of trade trade and sector (as opposed in the individual approach above), and therefore the only sources of variability correspond to the bilateral commodity-specific preference parameters and transport costs: b_{id}^{c}, τ_{id}^{c}, and τ_{nd}^{c} available for any trade flow. The extended specification therefore corresponds to the following equation:

Considering this estimation strategy results in the following specification:

$$\begin{aligned} \ln V_{id} &= \alpha_0 + \alpha_d + \alpha_i + \alpha^{c} + \beta_i^{c} \ln(1+\tau_{id}^{c}) + \beta_n^{c} \ln(1+\tau_{nd}^{c}) + \beta_1 \ln Dist._{id} + \\ &+ \beta_2 \ Bord.region_{nd} + \beta_3 \ Bord.country_{fd} + \beta_4 \ Adj.region_{nd} + \beta_5 \ Adj.country_{fd}, \\ &\quad i = n, f, \ \ c = 1, \dots, C. \end{aligned} \tag{12}$$

The corresponding econometric model, also estimated through PPML methods, identifies the commodity specific trade elasticities by using sectoral dummies capturing the interaction between the sector specific transport costs and the different administrative levels at which trade takes place, while controlling, as before, for the import and export specific fixed effects, but introducing commodity (sector) dummies, and, once again, an additional set of ancillary dummies capturing the different levels of administrative contiguity between regions and countries, as well as those controlling for border effects in the trade flows. Once again, the foreign and national elasticities of trade are, once again, identified from the parameters associated to the bilateral variation in transportation costs: i.e, $\sigma_F^{c} = 1 - \beta_i^{c}$ and $\sigma_N^{c} = 1 - (\beta_i^{c} + \beta_n^{c})$. Besides the advantage stated in the previous section regarding the possibility of jointly determining the significance of estimates for intranational and international trade, we can now establish this result with regards to the whole set of sectors, which adds robustness to the results that are individually obtained for each sector as the previous section.

3.3. Measuring trade elasticities with aggregated country to region.

The theoretical model developed in the second section and the resulting specifications presented above are based on region to region flows. It is implicitly assumed that consumers can eventually differentiate between goods sourced domestically, from regions belonging same country (internationally), or from regions situated in a different country. While it seems clear that consumers can differentiate between domestic, intranational and international goods, it can be argued that except for *connoisseurs* of varieties within a specific sector, ordinary individuals would not be normally able to tell apart varieties coming from foreign countries. To increase the awareness at international levels the so-called *appellations d'origine* in the food and beverages sector, or country trademarks in manufacturing such as “Made in Germany” have been promoted at different levels. However, in the general case, one may wonder to what extent trade elasticities would change if rather than estimating the previous specifications using region to region international flows, one aggregates trade flows at the country level from the perspective of the importing region. If the attained results do not change significantly, this would imply that the differentiation of products from the perspective of the consumers, regardless the regional variability in international flows, is limited. Also from an econometric perspective, an advantage of this approach is that by aggregating flows at the country level in origin, problems eventually related to small samples for some sectors and pairs of countries can be overcome. With high dimensional data (e.g., as in case of a large number of dummy regressors resulting in low degrees of freedom), standard regression estimators yield unstable coefficient estimates with inflated standard errors (Bühlmann and van de Geer, 2011), leading to reduced statistical power and erroneous conclusions regarding the significance of the trade elasticities in our case.

For both empirical and statistical reasons we run the same pair of regressions corresponding to the individual sector estimates of trade elasticities *à la* Hummels (2001) and Hertel *et al.* (2007)—eq. (11), and our proposal pooling trade data by sectors—eq. (12) , but aggregating international trade flows in by country of origin. Consequently, imports flows V_{fd} to region *d* are now aggregated for the *f* = 1,…,27 countries with which trade takes place, while the transportation costs associated to these flows τ_{fd}^{c} are averaged at the country level. As for the import flows and corresponding transportations costs from regions within the same country, these are kept as in the previous specifications, while the definition of adjacency (or contiguity) and border dummies is qualified to account for the specific case of adjacent regions belonging to the different countries: $Bord.reg.country$. This dummy variable is included in an attempt to capture the idiosyncratic nature of border regions belonging to different countries, and whose ties are being strengthened through the so-called euroregion structures (i.e., activities of Cross-Border Cooperation, CBC). A Euroregion can be very simply defined as a territorial unit formed by two contiguous sub-national units belonging to two separate states, which are involved in CBC activities (Perkmann, 2002; Durà et al., 2018). We contend that these integrating efforts may be successful in shaping the preferences of consumers towards products produced in neighboring border regions, and therefore report and comment on these values in the empirical section.

Having in mind this setting, the specification of the gravity equations underlying the calculation of country to region trade elasticities corresponds to eqs. (11) and (12), but changing the dependent and explanatory variables as discussed and including the $Bord.reg.country$ dummy variable to capture cross-border effects. This result will be presented and discussed in the empirical section under the heading “country to region” trade elasticities.

4. Data: Trade flows, generalized transport costs and control variables

The estimation of the trade elasticities model through equations (11) and (12) depending on the assumption of constant (perfectly competitive) or increasing return to scale (monopolistic competition) is done at the two digit level of the Statistical Classification of Products by Activity (CPA 2.1), corresponding to tradable goods, i.e., codes A, B, C, D and E, of the NACE Rev. 2 as presented in Appendix 1. This level of aggregation for trade flows corresponds to that available for the RHOMOLO GCE model, Lecca et al (2019). Agriculture (A) is the sector normally treated as perfectly competitive in trade models, while the rest of industries are the imperfectly competitive sectors. We consider 2013 as the reference year since this is the latest one available in the trade database.

The data used in the estimation can be grouped in three categories: 1) Trade flows (quantity and values); 2) generalized transportation costs and associated iceberg values, and 3) ancillary control (dummy) variables regarding contiguity (adjacency) so as to capture border effects when administrative boundaries are considered when segmenting trade flows into intraregional, interregional or international trade flows, as well as language and cultural proximity.

4.1.Trade flows

Trade data comes from the latest calculations of the EU REMO team and PBL Netherlands following the methodology proposed by Thissen et al. (2019). These authors estimate a probabilistic trade flows matrix to construct the inter-regional trade flows for all 267 (NUTS-2) EU regions. The methodology relies on 2013 national supply and use tables (SUTs), which are an update of the information of Eurostat SUTs, classified according to NACE Rev2, and corresponding to 2010.[12] A general discussion of the methods can be found in Mercenier et al. (2016). The Eurostat tables account for the distribution of re-exports over the origin and destination countries, ensure consistency in bilateral trade (i.e., import trade flows are consistent with export trade flows); and, finally, make certain that exports and imports of all regions add up to the national aggregates found in the country tables (i.e., top-down compatibility).

The estimation of bilateral trade flows among all EU regions corresponds to 2013. In a first step, inter-regional SUTs for 240 NUTS2 regions are estimated using the constrained quadratic minimization procedure by combining the regional Social Accounting Matrices and considering Thissen et al. (2019) data on inter-regional trade flows as priors.[13] In a second step, trade flows for the missing EU regions are estimated. The end result is a

---

[12] The detailed description of the NACE Rev2 classification can be found at Eurostat's Reference and Management of Nomenclatures (RAMON) site:
https://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC.

[13] For the specific optimizing function and set of restrictions see Thissen et al. (2019: 13-15).

regional trade matrix that is not only consistent with the regional SUTs, but also as close as possible to the main European transport data.

The specific sectors that we consider in our study corresponds to the two digit codes included in the A, B and C NACE classification, as presented in the Statistical Classification of Products by Activity in the European Union, Version 2.1, (CPA 2.1), appendix 1. For these sectors we have detailed trade information used for the estimating of the elasticities of substitution. In particular the imports (CIF) are used throughout the analysis as dependent variable of the gravity equations (11) and (12). This values are then used to calculate the ad valorem transport costs as the proportion of the transport cost over the price in origin. We return to this point later in section 4.3.

4.2. Generalized transportation costs

The calculation of the transport costs entering our econometric specification enhances existing approaches based on the optimization model that identifies the route that minimizes the road freight cost between an origin and a destination, taking into account the existing distance and time economic costs from a transport engineering and logistics perspective and the actual road network, see Combes and Lafourcade (2005), Zofío et al. (2014). Persyn et al (2019) employ this methodology to calculate a dataset of generalized transportation costs for the EU regions, GTCs. However, their GTCs do not differentiate by type vehicle or the commodity that is transported, neither take into account the degree of urbanization between the origin and destination (as in previous approaches they assume that all freight flows are carried by a reference 40 ton, 5 axle, articulated truck).

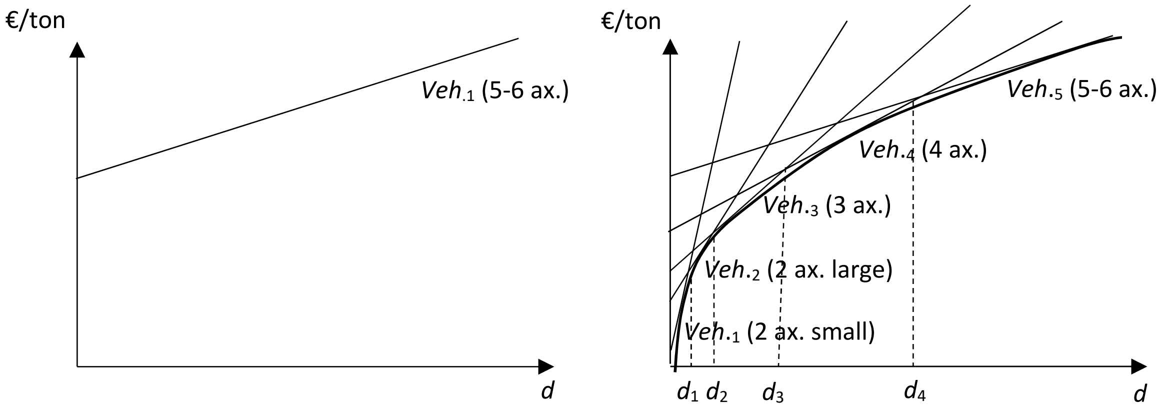
This differentiation is crucial in models estimating trade elasticities for several reasons. First, 'freight curves' determine the optimal vehicle size depending on the shipping distance between an origin i and destination j, d_{ij} (Jansson, 1982; McCann, 2001). Specifically, 'freight curves' identify the vehicle size that minimizes the cost per ton and per unit distance (e.g., €/ton/km). Second, coupled with distance, are the topological characteristics of the transportation network. The most salient feature is the road type, such as expressways, national or local roads, streets, etc., and the most limiting factor restricting the type of vehicle is whether the itinerary passes through urban areas, whose physical characteristics and regulations (based on risk or environmental concerns) only allow for small vehicles (i.e., light vehicles or 2 axle trucks). Hence, short distances corresponding to shipments within urban areas and conurbations (i.e., corresponding to a high proportion of intraregional trade that takes place within these configurations) are normally performed with this type of vehicles, while for intermediate distances medium size vehicles are preferred (i.e., 3-4 axle trucks), and, finally, longer distances are served with heavy duty vehicles (i.e., articulated trucks with 5-6 axles). Third, the type of vehicle employed for the shipment depends critically on the commodity transported (e.g., whether the cargo is dangerous, liquid or solid bulk, palletized, containerized, etc.). Both 'freight curves' determining the optimal size of the vehicle depending on distance, $s(d_{ij})$, human settlement patterns such as the degree of urbanization between the origin *i* and destintation *j*, u_{ij}, along with the type of commodity, *c*, restrict the vehicle specification that must be employed when establishing the distance and time economic costs underlying the GTC. This results in a specific selection of representative vehicles like the one we choose when calculating the GTCs. Consequently, vehicle specification, *v*, is a function of the previous variables: $v(s(d_{ij}),u_{ij},c)$.

4.2.1. 'Freight curves' and optimal vehicle size.

McCann (2001) relies on an inventory optimization approach to prove that under very general conditions the optimal size of a vehicle increases with the haulage distance and weight. In this regard the choice of optimal vehicle size based on distance hinges upon the balance between the structure of economic costs with respect to haulage distance, which are identified though transport engineering, and handling costs, determined through logistics. Haulage economic costs refers to the annual direct (based on distance and time) and ancillary indirect costs in which transportation firms incur when using a specific vehicle of their fleet. For illustrative purposes and taking the Spanish case as reference in 2018, the average annual cost of operating a heavy duty vehicle, HDV, corresponding to a 40 ton articulated truck with 5 axles and a 13.6 meter trailer 4 meters high—the typical 'workhorse' of the European road freight industry—is 127,646.89€/year. Handling economic costs refers to time costs associated to loading, unloading and docking operations (the latter including also the time spent in administrative paperwork upon arrival to and departure from the terminal). While haulage costs are variable by depending on distance, since the annual cost can be converted to cost per unit of distance by dividing by the amount of kilometers covered by the vehicle (e.g., for the HDV it is assumed that it travels 102.000 km per year fully loaded, resulting in 1.251€/km), handling costs are fixed and related to single legs (e.g. for the HDV, assuming that it is fully loaded with standard euro pallets, its takes about three hours and a half to complete the whole handling sequence corresponding to docking and loading and unloading logistics, Burdzik et al. (2014)).

The relationship between optimal vehicle size and distance is precisely driven by the trade-off between distance related costs expressed in euros per ton per km., that are lower the larger is the vehicle because it can carry a larger payload cargo (e.g., the maximum payload cargo that the HDV can carry is 25 tons), and handling operations whose time costs per ton are higher (as they take longer) the larger is the vehicle (as presented above). These can be clearly seen in Figures 1a and 1b. Figure 1a presents the 'cost line' associated to the HDV, identified as Vehicle 5 (*Veh.* 5), with a 5 or 6 axle configuration. For this vehicle the line represents the variable cost associated to distance, and whose slope is precisely 0.050€/ton/km, while the handling costs are the intercepts. Fig. 1b presents the 'cost lines' for vehicles increasing in size (as identified by the number of axles) from the smallest vehicle (light vehicle or small truck with 2 axles), *Veh.*1, to the largest one, $Veh._5$ (HDV with 5.-6 vehicles).

When comparing the cost per ton functions for the successive vehicles increasing in size, and the handling costs, the ‘freight curve’ naturally emerges as the envelopment from below of the alternative ‘cost lines’. This relationship rest upon a systematically negative relation between the size of an individual vehicle and its hauling and handling cot per ton.



Source: Jansson and Shneerson (1982: 226-227)

Figure 1a and 1b. Cost line for the HDV articulated vehicle (1a) and `freight curve’ (1b)

Based on transport engineering and logistics data this relationship is observed for the case of road freight transportation. Updating the economic costs data for the HDV presented in Zofío et al (2014: Table 1) for the HDV, and enlarging the database to include four additional vehicles decreasing in size (i.e., from the largest reference vehicle, *Veh.* 5, to the smallest vehicle, *Veh.* 1), as well as handling times from Burdzik et al. (2014), Table 1 presents the set of critical distance thresholds (d_1,..., d_4) that identify the distance at which each vehicle is optimal by minimizing the transportation cost. Each ‘cost line’ is defined by a fixed cost associated to handling operations that is decreasing in vehicle size, and variables costs corresponding to the hauling distance that are decreasing in vehicle size. The thresholds reported in the last column are calculated as the intersection points between the successive ‘cost lines’. The obtained results show that up to a distance of 10 km, the small vehicle is the optimal choice. The difference between two successive thresholds show the distance range in which a given size is optimal, i.e., that between the lower and upper thresholds. For example, the rigid vehicle with 3 axles, *Veh.*$_3$, is optimal in the 18 km range between 25 km and 43 km. Finally, aggregating consecutive thresholds yields the distance at which a given vehicle becomes optimal. For the HDV the cumulated distances show that it is the optimal vehicle choice for shipments longer that 150 km. Given these results, and the proximity of the distance thresholds, it seems unnecessary to consider all five types of vehicles when calculating the generalized transportation costs, thereby reducing the computing time necessary to perform the analysis (particularly when the information regarding optimal vehicle size is coupled with the degree of urbanization and type of committee as we show next). Consequently, in our analysis we consider three types of vehicles (shaded in gray in Table 1): the small vehicle that represents the preferred size up to 10 km, the intermediate rigid (3 axles) truck, which is optimal between 35km and 150 km, and the largest HDV, which is the vehicle of choice for shipments longer than 150 km.

Table 1: Distance thresholds for optimal vehicle sized depending on handling and hauling costs.[14]

	Maximum Payload (a)	Time costs (b)	Handling		Hauling	Distance
			(c)	(d) = (c)*(b)/(a)		
Vehicle	tons	€/hour	hours	€/ton	€/km/ton	Km
HDV (5 axl.)	25.0	30.4	3.5	4.3	0.050	72.0
Rigid (4 axl.)	22.3	27.7	3.2	3.9	0.058	43.0
Rigid (3 axl.)	16.0	24.9	2.1	3.3	0.073	25.0
Large (2 axl.)	9.5	22.7	0.9	2.2	0.114	10.0
Small (2 axl.)	6.0	21.0	0.4	1.3	0.206	

Notes: Own calculations.

To ease the comparison between the economic costs corresponding to each type of vehicle accounting for whether they are variable (depending on distance or time) or fixed, we present in Annex 2 their corresponding factors of proportionality with respect to the reference HDV considered by Persyn et al. (2019).[15]

4.2.2. Freight transport in urban areas.

As anticipated, besides the existence of optimal vehicles for alternative distances there are further constraints that limit the use of the above reference vehicles. These constraints refer to road infrastructure (in particular the urban grid or layout) and regulatory legislation (national, regional or city ordinances with respect to traffic congestion, safety, air pollution, etc.). The latter are intended to internalize the negative social and environmental impacts of urban freight transport. For this, reason when calculating the generalized transportation costs between an origin and a destination, it is necessary to combine vehicle optimality with

---

[14] Maximum payload, time cost and hauling cost are calculated based on information from the Observatory for Road Freight Road Transportation, MFOM (2018). Handling times for the HDV are reported by Burdzik et al. (2014). These include docking, loading and unloading operations. For the rest of vehicles loading and unloading times are calculated using the proportional rule given the capacity of the vehicles. For administrative and docking operations it is assumed that they do not apply to the small vehicle, while they are increasing in time the larger is the vehicle. For convenience in the GTCs calculations distance thresholds have been rounded to nearest whole (natural) number.

[15] Detailed data on actual monetary values by cost category for these vehicles as in Zofío et al (2014) are available upon request.

respect to distance and the reality of the geographical location in terms of their degree of urbanization. Specifically, city logistics and supply chain management make the small vehicles the only choice for the delivery of goods. There is a complex relationship between the spatial and functional structure of city logistics where the organization and density of land uses (i.e., degree of urbanization) interact with various forms of transport infrastructure, see Giuliano et al. (2019).

Although, cities present a variety of forms and levels of density, each associated with specific city logistics patterns, the available geographical information from the Global Human Settlement Layer (GHSL) project form the European Commission allows us to differentiate between three urbanization patterns.[16] The GHS model classifies human settlement according to certain rules of population and built-up density and contiguity of grid cells. Combining satellite information on the density of land use (built-up area) and census data, the GHSL method generates raster data of 1 km^2 resolution that differentiates between urban centers, urban clusters, and rural areas. A succession of grid cells presenting a population density larger than 1,500 inhabitant each, or more than 50% of built-up area, with a minimum total population of 50,000 individuals, is classified as an *urban center* (the main requirement is grid contiguity with 4-connectivity and allowing for gap filling). An *urban cluster* is a succession of cells totalizing more than 5,000 individuals, where each cell presents a populations density larger than 300 inhabitants. Finally, *rural areas* corresponds to a succession of inhabited grid cells without a population thresholds, with total population of less than 5,000 inhabitants.[17]

Combining the information of optimal vehicle size by distance, city logistics and degree of urbanization for distances in the range between 10km and 150km, the choice of representative vehicle corresponds to the origin-destination matrix presented in Table 2 (for distances shorter than 10 km and longer than 150 km, the preferred vehicles are the small 2 axle truck and the HDV, respectively):

---

[16] See Global Human Settlement (GHS) project: European Commission: https://ghsl.jrc.ec.europa.eu/data.php.

[17] Here, it is relevant that the GHS project uses grid cells to measure human settlement (built-up area) regardless of administrative boundaries, while census data includes a total count of individuals for administrative units varying widely in in size and shape, as well as population settlement and density within the areas. The GHSL method superimpose these two layers to create the new layer that also disregards administrative boundaries. This layer is segmented in grid cells of 1 km^2, and based on the population and land use thresholds, classifies the territory in the above categories. This constitutes the S-MOD module (settlement model).

Table 2. Representative vehicles combining optimal vehicle size, city logistics and urban patterns.

10 km < $d_{ij} \leq$ 35 km		destination j		
		Urb. Center	Urb. Cluster	Rural
origin i	Urb. Center	Small	Small	Small
	Urb. Cluster	Small	Small	Rigid
	Rural	Small	Rigid	Rigid

35 km < $d_{ij} \leq$ 150 km		destination j		
		Urb. Center	Urb. Cluster	Rural
origin i	Urb. Center	Rigid	Rigid	Rigid
	Urb. Cluster	Rigid	Rigid	HDV
	Rural	Rigid	HDV	HDV

Note: Small vehicle: 2 axles. Rigid vehicle: 3 axles. Heavy duty vehicle (HDV): 5 axles.

4.2.3. Economic costs by commodity

The last dimension in the calculation of the generalized transportation cost is to account for the type of commodity that is being transported. The reason is that the choice of vehicle depends on the commodity or, more generally, the physical characteristics of cargo that is transported in terms of weight and volume. Thus, the standard HDV is the vehicle of choice if the commodity can be transported in standard euro pallets (Burdzik et al., 2014). If perishable goods are transported (i.e., food) then it is necessary a temperature-controlled body made of insulated material and designed to carry temperature-sensitive (chilled or frozen) products. Then, if liquids, gases or powders (bulk cargo) are transported a tank fitted to a chassis is required. Other examples include the transportation of vehicles or containers that require ‘skeletal’ trailers. Modification of the above are also necessary in the case hazardous materials, wide loads, etc.

This variety of commodities results in substantial differences in economic costs across vehicles, and therefore, when calculating the GTCs associated to a given economic sector, where the cargo presents particular characteristics, one needs to control for the costs associated to the choice of vehicle required for transportation. How this is achieved can be easily exemplified for the case of sector 'C19’ in the Statistical Classification of Products by Activity in the European Union, Version 2.1, (CPA 2.1), corresponding to "Coke and refined petroleum products". The European Commission provides a matrix relating the CPA 2.1 to the Standard Goods Classification for Transport Statistics, 2007 (NST 2007), according to which shipments are classified in the European Freight Road Transportation survey (ERFT).

[18] This survey allows us to ultimately identify the type of cargo and associated vehicle. Reading the matrix of correspondences, sector 'C19' in the CPA 2.1 presents a one-to-one match with division 07 in the NST 2007: “Coke and refined petroleum products”. Hence tabulating the commodities transported data in the ERFT classified according to the NST 2007 and their associated type of cargo, one finds that for this particular sector the cargo corresponds mainly to liquid goods in bulk, 90.2% (of which 82.4% are dangerous—gasoline, gas, etc.—and the remaining 7.8% are non-dangerous), and solid goods in bulk, 9.8% (of which 5.9% are dangerous—solid or waxy refined petroleum products—and the rest, 3.9%, are not).[19] Using the ERFT survey for years 2011-2014, we are able to match the commodity transported and corresponding vehicle.

Appendix 3 presents the commodity factors that either increase or decrease the economic costs associated to each type of commodity taking as reference the standard HDV. For example, while cost of the HDV is 1.251€/km, that of a tanker increases to 1.590€/km (the annual costs in 2018 are 143,062.89€/year, and assuming that it covers 90.000km/year then the cost is 1.590€/km). Hence, the commodity factor between these two vehicles is 1.280 factor. Since the commodities belonging to a given NST 2007 classification are transported with a combination of vehicles (e.g., depending on whether dangerous and non-dangerous), the economic factors are the average of different vehicles weighted by the share of shipments transported by each type of vehicle.

4.2.4. Calculating the generalized transportation costs

Following Persyn et al. (2019), but allowing for the optimal type of vehicle depending on distance, urban characteristics, and commodity transported, $v(d,u,c)$, we denote by GTC_{ij}^{v} the generalized transport cost corresponding to the cheapest itinerary, I_{ij}^{v*}, among the set of possible routes $\mathbf{I}_{ij}^{v}$, of moving vehicle v between origin i and a destination j. The itineraries are comprised of different arcs a_c, with an associated set of physical and legal attributes (i.e., maximum legal speed), $\mathbf{x}_{a_c}$. The primary physical attributes of an arc are its distance, d_a, road type, r_a, and gradient (steepness), g_a. The arc speed, s_{a_c}, is derived from these properties (see Persyn et al. (2019;8-9). It is possible to determine the time it takes to cover it, $t_{a_c}^{t} = d_a^t / s_{a_c}^t$. As a result the physical characteristics of an arc are ultimately summarized by its associated distance and time variables: d_a^t and $t_{a_c}^t$.

---

[18] The detailed description of the CPA 2.1 and NST 2007, along with their concordance, tables can be found at Eurostat’s RAMON site: https://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC.

[19] In the European Freight Road Transportation survey, the specific goods-related variables (A3) used are the type of good and vehicle are A3_1 (Type of goods, NST 2007), A3_3 (Classification of dangerous goods) and A3_4 (Type of cargo). The matching between the type of cargo and the most suitable vehicle is comes from the observation provided by the Spanish Observatory of Freight Road Transportation.

The generalized transportation cost for a given commodity *c*, GTC^{c}_{ij}, corresponds to the solution to the following problem:

$$GTC^{c}_{ij} = f^{c} GTC^{v}_{ij} = f^{c} \min_{l_{ij} \in \mathbf{l}_{ij}} \left(DistC^{v*}_{ij} + TimeC^{v*}_{ij} \right) + Taxes^{v}_{ij} + Vignette^{v}_{ij} + Handling^{v}_{ij}, \tag{13}$$

where

$$\begin{aligned} DistC^{v}_{ij} &= \sum\nolimits_{a \in l_{ij}} \left(\sum\nolimits_{k} e^{vd}_{ak} f^{vd}_{ak} \right) d_a = \\ &= \sum\nolimits_{a \in l_{ij}} \left(fuel^{v}_{a} + toll^{cv}_{a} \right) d_a + \left(tireCS^{v} + maintCS^{v} \right) \left(fuel^{v}_{a} d_a \right), \end{aligned} \tag{14}$$

$$TimeC^{v}_{ij} = \sum\nolimits_{a \in l_{ij}} \left(\sum\nolimits_{k} e^{vt}_{ak} f^{vt}_{ak} \right) t_a = \sum\nolimits_{a \in l_{ij}} \left(1 + amortFinCS^{v}_{a} + insurCS^{v}_{a} + indCS \right) \left(t_a lab^{v}_{ij} \right). \tag{15}$$

Compared to Persyn et al. (2019: 3-4), the GTC in (13) has been adjusted by allowing for the choice of optimal type of vehicle depending on distance, urban characteristics and commodity transported *v*(*d, u, c*).[20] Thus, the original distance and time cost of the HDV considered by these authors, e^{d}_{ak} and e^{t}_{ak}, are modified by applying the individual vehicle factors corresponding to distance and time costs f^{vd}_{ak} and f^{vd}_{ak} (appendix 2), thereby obtaining the new costs at the arc level $e^{vd}_{ak} = e^{d}_{ak} f^{vd}_{ak}$ and $e^{vt}_{ak} = e^{t}_{ak} f^{vt}$. Unfortunately, from a practical perspective, and due to lack of data, it is impossible to control for commodity attributes at the arc level (e.g., such as legal traffic restrictions depending on hazardousness). Therefore, when calculating the minimum cost along the optimal itinerary only the information on distance and urban degree is considered, resulting in GTC^{v}_{ij}. For this reason, the final GTC associated to each commodity *c* as presented in (13), GTC^{c}_{ij}, is the result of multiplying this baseline GTC^{v}_{ij}, associated to the vehicle of choice, by the commodity factor f^{c}.

Following Persyn et al. (2019; 4-5) the calculation of a representative generalized transportation cost between any two regions is an aggregate measure of numerous GTC^{c}_{ij} between a random selection of origin and destination centroids, each drawn within a

---

[20] For a detailed discussion of each of the distance and time economic costs e^{d}_{ak} and e^{t}_{ak} (in EUR per km), see Persyn et al. (2019; 3-4). The main components of distance cost are fuel costs ($fuel_a$), which is computed as the fuel price (in EUR per liter) multiplied by the fuel consumption of the reference vehicle, and toll costs ($toll_a$) which are specific to each member state because of differences in nation-wide tolling policies (e.g., either through vignettes, or a country-wide electronic toll) or also per road-segment. The main time cost is the labor cost of the driver ($t_a lab_{ij}$). The hourly wage cost lab_{ij} from Eurostat is multiplied by the time (in hours) it takes to cross the arc. Labor costs correspond to the average wages at origin and destination. The remaining costs are proportional to the cost shares (*CS*) of these main components, based on the cost structures provided by the Spanish Observatory of Freight Road Transportations in 2018 (MFOM, 2018).

1kmx1km grid, where the number of centroids between grids depends on the population distribution in the regions. As a result we calculate the GTC between two regions *o* and *d* as the arithmetic mean of the GTC between the *I* centroids belonging to region *o* indexed by *i*=1,…,*I*, and the *J* centroids belonging to region *d* indexed by *j*=1,…,*J*. The final inter-regional GTC equals for a given commodity *c* corresponds to: $GTC_{od}^{c}=\frac{1}{IJ}\left(\sum_{i}\sum_{j}GTC_{ij}^{c}\right)$. In the above calculations it is possible to identify the type of vehicle that is used at each centroid pair levels and, by taking the average, the corresponding GTC associated to each one of them.

4.3. Ad valorem transportation costs

We are now ready to present the calculations of the iceberg transportations costs that are included in the econometric specification obtained from the trade model. The iceberg transportation cost defines in the trade literature as a "wasteful ad valorem" penalty, meaning that they are assumed to be proportional to the value of the good, with a constant fraction "melting" away, and implying that some extra proportion needs to be shipped for the intended quantity to arrive at destination.

In practice we have seen that real (generalized) transport costs depend on a number of characteristics related to the choice of vehicle. Following Persyn et al (2019), we can take advantage of the GTC calculations when defining the ad valorem transportation cost between any two regions for each trade sector. Considering that we can match trade flows classified according to the CPA 2.1 with their corresponding generalized transport cost following the NST 2007 classification; i.e., to each sector *s* =1,…,*S* we associate the GTC for commodity *c* = 1,…, *C* (see appendix 1), we define the iceberg transport cost τ_{od}^{s} between regions *o* and *d*, for each sector *s*, as follows:

$$\tau_{od}^{s}=\frac{F_{od}^{s}\sum_{v=1}^{3}\left(\frac{s_{od}^{v}}{L_{od}^{v}}\right)GTC_{od}^{v}}{V_{od}^{s}}=\frac{\sum_{v=1}^{3}\left(\frac{s_{od}^{v}}{L_{od}^{v}}\right)GTC_{od}^{v}}{P_{o}^{s}},\ s_{od}^{s}\geq 0,\ \sum_{v=1}^{3}s_{od}^{s}=1, \tag{16}$$

where F_{od}^{s} (tons) and V_{od}^{s} (€) are the quantity and value of the trade flows in origin; GTC_{od}^{v} (€/veh.) is the generalized transport cost for each vehicle size, calculated as in (13); s_{od}^{v} are the transportation shares of each vehicle in the bilateral shipments; and, finally, L_{od}^{v} (tons/veh.) is the average load of the shipments. The information on vehicles shares s_{od}^{v} and average loads L_{od}^{v} is obtained from the European Fright Road Transportation (ERFT) survey by segmenting the sectoral shipments according to the maximum permissible laden weigh, and calculating the ratio of actual payload to maximum payload. [21] The numerator in

---

[21] With regard to Persyn et al (2019: 5), the current iceberg transport cost specification accounts for the variability in the type of vehicle used in the shipments (rather than relying on the single HDV), which should

expression (16) calculates the number of vehicles necessary to ship the quantity F_{od}^{s} according to current distribution of vehicles, by multiplying the number or required vehicles by their cost. Subsequently, the transport cost is related to the value of the shipments, yielding the ad valorem value.

4.4. Control variables

Finally, ancillary variables, geographical distances, adjacency, common language are idiosyncratic characteristics that are taken into account in trade for each pair of regions, as they may represent relevant enablers or barriers to bilateral trade. Distances between regions as well as information about contiguity of regions and countries are computing using the Geodata on Administrative Units provided by Eurostat GISCO. Also, we use geodesic distances, calculated by computing the distance between the physical centroids of the regions.

5. Results

5.1. Individual (intra)national and foreign (international) trade elasticities by sector

We present the results obtained from estimating the individual sector foreign and (intra)national trade elasticities using region to region data—eq. (11)—and country to region data (as described in section 3.3) in Tables 3 and 4, respectively. Focusing first on the former level of aggregation, the classic estimation strategy that estimates both trade elasticities by individual sectors cannot capture the two levels of substitutability between goods produced in regions within the same country, and those coming from regions situated in a different country. Based on the statistical significance of the parameters, the third row of the table presents whether foreign trade elasticities are greater, equal, or smaller than national trade elasticities.

It can be observed that only in 2 out of the 19 sectors both trade elasticities are significant; i.e., C24 (‘Basic metals’) with σ_F^{C24} = 4.141 and σ_N^{C24} = 6.231, and C27 (‘Electrical equipment’) with σ_F^{C27} = 6.649 and σ_N^{C27} = 26.721. In these cases, the hypothesis that

---

capture the reality of transport costs in short distances, and allows for different average loads across origin-destination pairs. It is well known that the load factor (asymmetrically) depends on *od* and the type of good transported (e.g., automobile-carrying vehicles, just like tanker trucks will by their very nature complete the majority of their trips empty– between 45% and 50%. In the case of the standard HDV, with vehicles carrying all types of cargo, hauling them to many locations, productivity in terms of the average load will depend on several factors. For example, in countries with predominantly national transport (like France or Germany) with shorter journeys and a strong commercial presence on the territory, the empty run rate will be low: less than 20% in Germany, and 13% in France. To be noted that in Spain, where transport increasingly tends to focus on the national territory, the empty run rate is also low, approximately 15%. In addition, in these countries, consolidations are widely applied, which allow to increase the vehicles’ load factor. More than half of loads are made up of consolidations (https://teleroute.com/en-en/).

regional elasticities should be higher than foreign elasticities based on the assumption that intranational trade is more sensitive to price variations than international trade as a result of higher substitutability is confirmed. This result may be consequence of the fact that intranational trade faces fewer non-price trade restrictions than international trade (even within single markets like the EU) while, at the same time, the goods (varieties) produced in regions within the same country exhibit higher homogeneity. However, for the majority of the remaining sectors, either the parameter associated to the foreign or the national trade elasticity is not statistically significant, and therefore it is concluded that there is no difference between the two: $\sigma_F^c = \sigma_N^c$. That is, from the perspective of consumers, trade within the EU would not essentially differ depending on whether the region of origin situates within the same country or in a different country.

Trade elasticities range between that corresponding to 'Mining and quarrying' (sector B, σ_F^B = 1.651) and 'Fabricated metal products except machinery and equipment' (Sector C25, σ_F^{C25} = 44.777). In this case both values are recovered from the international trade flows. Out of the 14 sectors where foreign and national trade elasticities coincide, 6 elasticities are recovered from international trade flows, while the remaining 8 are recovered from (intra)national or regional trade flows. As previously mentioned only two sectors present different elasticities with the national value being greater than the foreign value. Finally, one sector, C21 ('Basic pharmaceutical products and pharmaceutical preparations') does not yield significant results, while two sectors present *negative* trade elasticities resulting from the positive relation between transportation costs and trade flows. These two sectors are C29 ('Motor vehicles trailers and semi-trailers'), and C30 ('Other transport equipment'). Inspecting the relationship between imports and transportation cost we find that for these two sectors, the density of trade is increasing in the transportation cost (but not distance), suggesting that it is the other elements of the iceberg `tau´ definition (16) what drives this relationship (i.e., average loads and/or units prices).

As for the rest of the control variables they present the expected sign and significance. The estimated parameter for the log of distance (ln*Dist*) is always negative and generally significant, while those associated to the regional and national border dummies capturing trade within regions and/or countries ($Bord.region_{id}$ and $Bord.country_{id}$), as well as geographical adjacency ($Adj.region_{id}$ and $Adj.country_{id}$), are also in general positive and significant.

Turning now to the country to region results presented in Table 4, we observe that aggregating the trade flows (and averaging transportation costs and distances accordingly) results in the loss of statistical significance in more than half of the sectors (10 out of 19). Although the correlation with the region to region trade elasticities for the remaining sectors is rather high (and normally identified from the same set of flows; i.e., intranational or international), we conclude that using the standard individual sector specifications, trade data at the EU country level does not allow the identification of the trade elasticities. An interesting exception are now sectors C29, ‘Motor vehicles trailers and semi-trailers’, and C30, ‘Other transport equipment’, whose values are unreliable when region to region flows are considered, while in this alternative setting, sensible magnitudes seem to be obtained. As for the control variables they also present sensible signs and significance, while the new (additional) dummy variable, controlling for the fact that trade takes place between neighboring regions situated in different countries, is negative and not significant in half of the sectors.

Table 3. International and intranational elasticities of trade (individual sectors. Region to region).

Variable \ Sector	A01	A02	B	C10-C12	C13-C15	C17	C19	C20	C21	C22
$\sigma_F = 1 - \beta_i$	–	–	2.642**	–	14.480*	–	–	2.465*	–	–
$\sigma_N = 1 - (\beta_i + \beta_N)$	7.352***	5.105***	–	25.703***	–	20.992***	8.691***	–	–	19.318*
σ_F >=< σ_N	=	=	=	=	=	=	=	=	–	=
β_ι	0.173	1.215	-1.642**	-1.349	-13.480*	-2.154	0.287	-1.465*	0.201	-5.576
	(1.037)	(0.748)	0.693)	(2.210)	(7.309)	(1.500)	(1.144)	(0.792)	(.)	(4.097)
β_N	-6.352***	-4.105***	-0.651	-24.703***	7.991	-19.992***	-7.691***	-1.527	-1.256	-18.318*
	(2.026)	(1.127)	(0.795)	(7.571)	(8.482)	(3.781)	(2.533)	(1.105)	(.)	(10.203)
ln*Dist.*	-0.312*	-0.367	-0.297***	-0.273*	-0.382***	-0.147	-0.322***	-0.016	-0.621	-0.204***
	(0.179)	(0.229)	(0.074)	(0.159)	(0.095)	(0.119)	(0.068)	(0.115)	(.)	(0.062)
Border.Reg	2.466***	2.028***	2.098***	2.077***	0.491	2.598***	1.243***	0.716**	-0.085	0.540***
	(0.605)	(0.582)	(0.252)	(0.391)	(0.329)	(0.412)	(0.230)	(0.323)	(.)	(0.205)
Border.Country	3.557***	7.133***	3.287***	4.484***	3.373***	4.015***	4.359***	3.541***	3.611	3.076***
	(0.341)	(0.608)	(0.719)	(0.587)	(1.039)	(0.597)	(0.332)	(0.281)	(.)	(0.232)
Adj.Region	1.007***	0.978***	1.032***	0.829***	0.288***	1.052***	0.562***	0.347***	0.123	0.246***
	(0.154)	(0.287)	(0.249)	(0.136)	(0.104)	(0.235)	(0.125)	(0.089)	(.)	(0.074)
Adj.Country	-1.189***	1.026*	-0.846	-0.592	0.063	(0.458	-0.389	0.706**	-0.184	0.766***
	(0.333)	(0.564)	(0.715)	(0.581)	(0.738)	(0.374)	(0.336)	(0.309)	(.)	(0.175)
Observations	69,179	68,124	69,172	69,182	69,193	69,437	69,169	69,222	65,069	69,250

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * p<0.10, ** p<0.05,*** p<0.01

Variable \ Sector	C23	C24	C25	C26	C27	C28	C29	C30	C31_C32
$\sigma_F = 1 - \beta_i$	–	4.141***	44.777***	4.438**	6.649***	–	–	18.213***	1.928***
$\sigma_N \underset{<}{\geq} 1 - (\beta_i + \beta_N)$	10.263***	6.231*	–	–	26.721***	7.267*	(-5.075)**	(-4.862)***	–
σ_F >=< σ_N	=	=	=	=	<	=	=	>	=
β_ι	-1.774	-3.141***	-43.777***	-3.438**	-5.649***	0.152	-1.394	-17.213***	-0.928***
	(1.222)	(0.947)	12.804)	(1.395)	(1.512)	(1.101)	(0.866)	(4.154)	(0.272)
β_N	-9.263***	-2.090*	4.541	-4.579	-20.072***	-6.267*	6.075**	23.075***	-0.785
	(3.515)	(1.270)	(20.260)	(5.427)	(7.515)	(3.453)	(2.700)	(7.985)	(0.696)
ln*Dist.*	-0.181	-0.062	-0.406***	-0.333***	-0.212***	-0.320***	-0.400***	-0.367***	-0.331***
	(0.180)	(0.102)	(0.119)	(0.062)	(0.072)	(0.077)	(0.091	(0.085)	(0.062)
Border.Reg	2.619***	0.907**	1.448***	0.751**	2.423***	0.887***	1.412***	1.574***	1.129***
	(0.463	(0.362)	(0.355)	(0.312)	(0.312)	(0.219)	(0.255)	(0.341)	(0.235)
Border.Country	3.659***	3.181***	3.632***	2.266***	1.426***	3.247***	0.882**	1.572***	2.632***
	(0.467)	(0.353)	(0.292)	(0.266)	(0.353)	(0.207)	(0.377)	(0.260)	(0.251)
Adj.Region	1.135***	0.320***	0.741***	0.291**	0.991***	0.428***	0.684***	0.588***	0.566***
	(0.188)	(0.121)	(0.118)	(0.128)	(0.164)	(0.080)	(0.164)	(0.180)	(0.106)
Adj.Country	0.863***	0.921***	0.298**	0.417***	0.488***	0.492***	0.422**	(0.046	0.825***
	(0.234)	(0.254)	(0.151)	(0.161)	(0.122)	(0.133)	(0.177)	(0.246)	(0.138)
Observations	69,696	69,486	69,185	69,224	69,180	68,981	69,195	69,253	68,167

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * p<0.10, ** p<0.05,*** p<0.01

Table 4. International and intranational elasticities of trade (Individual sectors. Country to region).

Sector / Variable	A01	A02	B	C10-C12	C13-C15	C17	C19	C20	C21	C22
$\sigma_F = 1 - \beta_i$	–	–	–	–	–	–	–	–	–	–
$\sigma_N = 1 - (\beta_i + \beta_N)$	–	–	–	19.770***	–	11.978***	–	–	–	–
σ_F >=< σ_N	–	–	–	=	–	=	–	–	–	–
β_t	3.370	-3.420	1.895	-5.433	15.991	-5.939	1.312	-0.969	-1.153	0.843
	(.)	(2.320)	(.)	(8.774)	(13.085)	(3.709)	(.)	(0.956)	(0.946)	(4.269)
β_N	-8.298	2.144	-4.761	-18.770**	-15.832	-10.978**	-6.269	-0.333	1.157	-8.838
	(.)	(2.019)	(.)	(8.542)	(13.432)	(5.344)	(.)	(1.251)	(1.111)	(10.669)
ln*Dist.*	-0.771	-0.562**	-0.398	-0.512***	-0.522***	-0.564***	-0.483	-0.295***	-0.705***	-0.406***
	(.)	(0.250)	(.)	(0.178)	(0.119)	(0.119)	(.)	(0.048)	(0.078)	(0.040)
Border.Reg	19.424	8.260***	8.974	8.673***	11.123***	3.508***	8.230	6.525***	6.895***	6.069***
	(.)	(1.362)	(.)	(0.777)	(1.462)	(1.266)	(.)	(0.607)	(0.807)	(0.286)
Border.Country	18.459	6.198***	8.280	7.529***	11.084***	1.900	7.429	6.533***	6.830***	6.083***
	(.)	(1.479)	(.)	(0.977)	(1.518)	(1.362)	(.)	(0.619)	(0.806)	(0.308)
Adj.Region	0.672	0.892***	1.024	0.672***	0.232**	1.171***	0.492	0.190***	0.173*	0.087
	(.)	(0.337)	(.)	(0.171)	(0.116)	(0.236)	(.)	(0.062)	(0.094)	(0.057)
Adj.Country	-0.905	0.284	0.257	-1.252*	-0.079	-0.546	-1.129	-0.166	-0.798**	0.498***
	(.)	(0.559)	(.)	(0.657)	(0.638)	(0.398)	(.)	(0.330)	(0.394)	(0.111)
Border.Reg.Count.	-0.434	-0.930	0.022	-0.494	0.292	-0.459*	0.431	0.024	-0.035	-0.129
	(.)	(0.605)	(.)	(0.502)	(0.266)	(0.253)	(.)	(0.247)	(0.230)	(0.102)
Observations	12,154	12,026	12,336	12,343	12,352	11,311	11,794	12,122	11,319	12,113

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Sector / Variable	C23	C24	C25	C26	C27	C28	C29	C30	C31_C32
$\sigma_F = 1 - \beta_i$	–	3.276*	141.832***	–	8.066**	–	6.509*	15.376*	2.739*
σ_N $1 - (\beta_i + \beta_N)$	19.252***	–	12.568*	–	–	–	(-7.250)***	–	–
σ_F >=< σ_N	=	–	>	–	=	–	<	=	=
β_t	9.408	-2.276*	-140.832**	-2.140	-7.066**	-2.482	-5.509*	-14.376*	-1.739**
	(5.774)	(1.327)	(66.219)	(3.159)	(3.482)	(3.930)	(3.015)	(8.639)	(0.678)
β_N	-18.252***	1.073	129.264*	-7.326	-17.553	-3.731	13.759***	22.482	1.004
	(6.026)	(2.309)	(74.491)	(4.640)	(13.298)	(4.671)	(3.320)	(14.313)	(0.663)
ln*Dist.*	-0.708***	-0.465***	-0.677***	-0.465***	-0.534***	-0.522***	-0.753***	-0.651***	-0.564***
	(0.103)	(0.081)	(0.054)	(0.046)	(0.076)	(0.044)	(0.124)	(0.063)	(0.080)
Border.Reg	8.648***	6.974***	6.541***	7.785***	3.877***	5.149***	4.966***	4.970***	4.371***
	(0.626)	(0.347)	(0.421)	(1.106)	(0.607)	(0.838)	(1.167)	(0.599)	(0.578)
Border.Country	7.729***	7.015***	5.472***	7.606***	2.343***	4.905***	4.231***	4.153***	3.623***
	(1.009)	(0.393)	(0.619)	(1.163)	(0.765)	(0.888)	(1.192)	(0.674)	(0.679)
Adj.Region	0.845***	0.150	0.568***	0.167	0.920***	0.227**	0.468**	0.289**	0.524***
	(0.273)	(0.155)	(0.150)	(0.123)	(0.213)	(0.093)	(0.197)	(0.142)	(0.127)
Adj.Country	1.369***	0.900***	-0.115	0.318***	0.253**	0.413***	0.199	0.484***	0.562***
	(0.428)	(0.158)	(0.289)	(0.120)	(0.123)	(0.148)	(0.207)	(0.182)	(0.133)
Border.Reg.Count.	0.290	-0.254**	-0.535*	-0.174*	-0.311***	-0.352***	-0.418***	-0.421***	-0.348***
	(0.282)	(0.129)	(0.306)	(0.094)	(0.106)	(0.135)	(0.129)	(0.131)	(0.122)
Observations	12,340	12,236	12,341	11,872	12,337	11,767	12,089	12,296	11,786

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

5.2. Pooled intranational and international trade data by sector

The standard method of estimating trade elasticities sector by sector can be compared to that of pooling the trade data as presented in eq. (12). This specification introduces the interaction between the transportation costs and their corresponding sectors to identify foreign elasticities of trade, and again with a dummy variable capturing whether the trade flow takes place between regions within a country to identify national elasticities of trade. As opposed to eq. (11) the advantage of this specification in that it yields a single value for the common variables that control for distance, border and adjacency effects, while allowing for a larger number of observations. By substantially increasing degrees of freedom, the confidence on the estimates improves by reducing their margin of error and allowing for reliable comparisons of trade elasticities across sectors. Table 5 reports the results relying of region to region data while Table 6 reports country to region results.

As in the previous tables the third row summarizes whether the foreign elasticity of trade differs from its national counterpart based on the significance of their associated parameters. Under this specification we now realize that foreign and national trade elasticities are indeed statistically different for most of the sectors. In all but 2 of the 19 sectors trade elasticities differ, with national elasticities being greater that foreign elasticities in 14 sectors. Only in 3 sectors the opposite is observed: C19, 'Coke and refined petroleum products', C21, 'Basic pharmaceutical products and pharmaceutical preparations' (whose value could not be recovered with the standard sector by sector estimation), and C25, 'Fabricated metal products except machinery and equipment'. As for sector A02, 'Products of forestry logging and related services', and C20, 'Chemicals and chemical products', foreign and national elasticities coincide, with its value been driven by international trade flows. As opposed to the standard approach of Hummels (2001) and Hertel et al. (2007) previously used, we remark that the estimation strategy corresponding to eq. (12), which provides better accuracy and test statistics, seems to be the appropriate one when recovering the two levels of trade elasticities corresponding to foreign and national goods. Now all parameters exhibit the expected sign and statistical significance (except for one sector: C10-C12, 'Food products beverages and tobacco products' whose sign is positive, resulting in a negative elasticity). Ultimately, the former approach does not allow for their identification, casting doubts about the applicability of these estimation method in single markets where tariffs are unavailable as the key identification variable.

The results obtained from aggregating regional imports by country of origin are presented in Table 6. Comparing the goodness of fit of these results with those obtained when estimating trade elasticities by sectors individually (Table 4), we observe that they are statistically significant for most of the sectors (while many were missing in the individual sector estimations). Also, the values of the elasticities correlate to a high extent with those obtained with region to region data (Table 5): $\rho(\sigma_F^{Reg.to.Reg.}, \sigma_F^{Country.to.Reg.})$ = 0.978, and $\rho(\sigma_N^{Reg.to.Reg.}, \sigma_N^{Country.to.Reg.})$ = 0.808. Relying on this aggregation we are able to recover the foreign trade elasticity for sector C13-C15, but again sector C21 dos not provide significant estimates for any of the trade elasticities. However, only 7 (8) sectors exhibit national elasticities greater (smaller) that their foreign elasticities (as opposed to the 14 (3) sectors in Table 5). Hence, aggregating regional trade flows by country of origin reverses the conclusions regarding the direction of the inequality between foreign and national trade elasticities, showing that the geographical level of analysis has relevant consequence in the magnitude of the trade elasticity estimates.

5.2.1. Foreign (international) elasticities of trade

Foreign elasticities of trade relying on region to region data (Table 5) range from the minimum observed in sector C31_32, 'Furniture and other manufactured goods', $\sigma_F^{C31\_32}$ = 2.106 and the maximum observed in sector C25, 'Fabricated metal products except machinery and equipment', σ_F^{C25} = 187.190. Nevertheless, this latter value represents an extreme as it is one order magnitude (ten times) larger than the next value corresponding to sector C30, 'Other transport equipment', σ_F^{C30} = 19.870. For comparison purposes national elasticities range between σ_N^{C21} = 1.779 (C21, 'Basic pharmaceutical products and pharmaceutical preparations') and σ_N^{C25} = 134.035 (again for sector C25). In general, foreign trade elasticities are smaller than 10, with their corresponding national counterparts more than doubling their value. These values of foreign elasticities of trade are in line with those reported in the literature relying on trade flows from projects such as GTAP (World data), Michigan (US), USAGE (US) and MONASH (Australia), using tariffs as identification variable, and time series or cross-sectional analyses as econometric approaches. Comparing our results to those reported in Table 1 by Hilberry and Hummels (2013; 1,221) for multicountry (including some EU countries or the block as a whole) and single-country models, the range of elasticities is [0.9, 34.4]. As the level of sectoral aggregation is similar to ours, we confirm that our estimates are in line with those provided by previous studies (see also Table 1 in Hertel et al., 2007).

There are also a few studies reporting Armington elasticities between EU countries such as Németh et al. (2011), Welsch (2008) and Olekseyuk and Schürenberg-Frosch (2016). The former authors use the so-called GEM-E3 model aimed at capturing the interactions between economy, energy and the environment in a general equilibrium modelling framework.[22] Using the European version of the model they report short and long-term Armington elasticities which are estimated relying on a panel data analysis econometric framework that uses dynamic adjustments. The dataset covers yearly data for the 1995–2005 period and the range of elasticities between domestically produced and imported

---

[22] For details on the model see https://ec.europa.eu/jrc/en/gem-e3/model.

goods (home-foreign goods) for seven energy-intensive sectors is [0.6; 1.7].[23] These values are particularly low in light of the elasticities reported by Hilberry and Hummels (2013) and our own estimates, whose minimum value, $\sigma_F^{C31\_32}$ = 2.106, is above their upper bound.[24] On their part Olekseyuk and Schürenberg-Frosch (2016) estimate country-specific Armington trade elasticities using trade data between eight EU countries for the manufacturing sectors. They use a panel data set constructed from the STAN-OECD and EUROSTAT's PRODCOM databases (for different ISIC classifications) covering the period 1995-2011. As for the estimation methods they consider single-sector co-integration time series analysis and, when the sample size is insufficient to identify the elasticities, turn to a fixed effects panel data model that coincides with our approach. By pooling the data for neighboring sectors they are capable of increasing the number of observations (i.e., the number of degrees of freedom) and, thus, the accuracy of the results and the test statistics. Regarding the use of time series, they adopt co-integration methods because their estimates show that for most countries both the price and quantity ratio series are non-stationary, but integrated of order one or two. Their range of elasticities corresponding to the manufacturing sectors considered in their study is [0.30; 3.67]. As for the panel data results, their pooled fixed effects estimations yield trade elasticities in the range [0.32; 2.43]. Focusing on the goodness of fit we highlight that, compared to the above studies, the number of sectors for which trade elasticities exhibit the right sign and are statistically significant is greater in our case (all but the foreign elasticity of sector C13-C15), thereby providing a complete set of results for all tradable sectors: agriculture and fishing, mining and quarrying and manufacturing. Altogether, we conclude that our foreign elasticities of trade are above those previously estimated for EU countries only, but in the range of those obtained in international studies including countries of several world regions.

As before, we now compare the region to region results reported in Table 5 to their country-to region counterparts in Table 6. Focusing on descriptive statistics and dropping the values that are not significant, average foreign trade elasticity considering region to region data is 17.941 (24.929 with country to region data), with a standard deviation of 42.537 (41.567) and median value of 7.846 (10.783). Hence, although the values correlate to a large extent (as shown above), aggregating trade flows geographically result in a light increase in the value of the foreign trade elasticity.

5.2.2 National (regional) elasticities of trade

As for the second level of trade, national (regional) elasticities of trade within the EU are new to the literature. Our results indicate considerable variability across sectors. Besides

---

[23] Previously, and focusing also on energy intensive sectors Welsch (2008) estimates elasticities for four European countries and 17 sectors with values ranging between 0.04 and 3.68.

[24] Németh et al. (2011) also report the trade elasticity between imports origination from two different countries (foreign-foreign) rather than between foreign and domestic (home-foreign) goods. These range between [0.8; 2.8].

the spread already presented ranging between σ_N^{C21} = 1.779 and σ_N^{C25} = 134.035, the majority of sectors (12 out of 19) exceed the value of 10. Excluding the values of the two sectors that are not statistically significant, average national trade elasticity is 26.929, with a standard deviation of 31.780 and median value of 18.510. Generally, it is observed that the smaller national trade elasticities correspond to sectors with relative low value added and/or producing relatively homogenous varieties. Beyond the lowest value σ_N^{C21} = 1.779, corresponding to sector C21 ('Basic pharmaceutical products and pharmaceutical preparations'), agriculture and fishing (A01), mining and quarrying (B), coke and refined products (C19), and basic metals (C24) exhibit elasticities below 10. Only sector C31-C32, 'Furniture and other manufactured goods', escapes this general characterization. On the contrary, values of trade elasticities above 10 correspond in general to sectors producing goods with higher value added and heterogeneous characteristics. Besides the largest value σ_N^{C25} = 134.035, all equipment related goods (computer, C26; electronic, C27; machinery, C28; and transport, C30), as well as motor vehicles (C29), and some chemical products (rubber and plastic, C22), show elasticities close to or above 25, thereby doubling the average elasticity for the previously listed low value added and homogenous sectors.

Although it was not previously discussed, this sectoral evaluation of the values of national trade elasticities applies straightly to their (lower valued) foreign trade counterparts since both series highly correlate: $\rho(\sigma_F^c, \sigma_N^c)$ = 0.900. Although counterintuitive, the fact that both foreign and national trade elasticities are lower for low-value added and homogenous goods and higher for high value added and heterogeneous goods is in line with the results reported in Olekseyuk and Schürenberg-Frosch (2016). Following these authors we recall a likely explanation proposed by Saito (2004) who presents evidence on why higher elasticities are observed in sectors trading mainly intermediate inputs, as it is the case for those listed above.

Even if for the European case there are no precedents in the estimation of national (regional) elasticities of trade, Bilgic et al. (2002) have estimated 'regional' trade elasticities among US states and compared them to their international (foreign) counterparts. To the extent that the US represents a single market area comparable to that of the UE, and therefore only transportation costs are available to identify elasticities, it is worth comparing our results with theirs.[25] Their estimates situate within the previous range [0.45; 2.80], and comparing them with those reported in previous literature at the international level, they conclude that regional (national) elasticities and greater than international (foreign) elasticities, as it is also our case. However, because they do not set up a comprehensive three-level model that allows to jointly estimate foreign and national elasticities as the one introduced in section 2, the actual magnitudes between both sets of

---

[25] These authors also review estimates from seven US regional CGE studies, all corresponding to gravity equation that are derived from a CES theoretical framework. The elasticities obtained in those studies for tradable goods sectors (i.e., excluding services) range between [1.50; 3.50].

results are not directly comparable (nor can it be determined if their differences are statistically significant) because they are the result of using different data samples, time periods and econometric specifications (regardless of how marginal the differences are).

The effect of aggregating import flows by country of origin on the estimates of national (regional) trade elasticities, when compared to their region to region counterparts, is much higher than with respect to foreign elasticities. While for foreign trade elasticities results remain numerically the same, aggregating import flows results in a noticeable reduction in the value of national trade elasticities. To the point already noted that the inequalities between foreign and national elasticities reverses: larger in favor of the latter with region to region data and the opposite with country to region data. Indeed, average national trade elasticity with region to region data is 26.929 (13.708 with country to region data), with a standard deviation of 18.510 (11.905) and a median value of 31.780 (9.084). Hence, although the values correlate to a large extent (recalling from above, $\rho(\sigma_N^{Reg.to.Reg.}, \sigma_N^{Country.to.Reg.}) = 0.808$), aggregating trade flows geographically changes one of the main conclusions of the study. This prompt us to advice caution to other researchers interested in calculating foreign and national trade elasticities simultaneously, as the geographical level of aggregation is critical. That is, one of the reasons to use alternative levels of geographical aggregation is to perform robustness checks and, in this case, using country-to-region data results in opposite conclusions.

Table 5. International and intranational elasticities of trade (pool sectors. Region to region).

Variable \ Sector	A01	A02	B	C10-C12	C13-C15	C17	C19	C20	C21	C22
$\sigma_F = 1 - \beta_i$	7.122***	6.633***	3.541***	17.104***	(-4.793)***	8.569***	9.037***	2.179***	2.181***	14.442**
$\sigma_N = 1 - (\beta_i + \beta_N)$	9.293***	–	4.085***	26.455***	16.800***	18.51***	6.951***	–	1.779**	31.475**
$\sigma_F \gtrless \sigma_N$	<	=	<	<	<	<	>	=	>	<
β_ι	-6.122***	-5.633***	-2.541***	-16.104***	5.793***	-7.569***	-8.037***	-1.179***	-1.181***	-
	(0.418)	(0.531)	(0.208)	(0.771)	(0.764)	(0.408)	(0.449)	(0.189)	(0.150)	(1.439)
β_N	-2.171***	0.428	-0.544***	-9.355***	-21.593***	-9.941***	2.086***	-0.607	0.402**	-
	(0.835)	(0.506)	(0.194)	(1.325)	(1.549)	(0.715)	(0.711)	(0.448)	(0.196)	(4.176)
ln*Dist.*	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Border.Reg	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***
	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)
Border.Country	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Adj.Region	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Adj.Country	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Observations	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * p<0.10, ** p<0.05,*** p<0.01

Variable \ Sector	C23	C24	C25	C26	C27	C28	C29	C30	C31_C32
$\sigma_F = 1 - \beta_i$	9.314***	4.774***	187.190***	3.323***	8.573***	12.700***	4.373***	19.870***	2.106***
$\sigma_N = 1 - (\beta_i + \beta_N)$	14.616***	9.355***	134.035***	23.318***	61.863***	23.927***	26.867***	45.578***	2.890***
$\sigma_F \gtrless \sigma_N$	<	<	>	<	<	<	<	<	<
β_ι	-8.314***	-3.774***	-186.190***	-2.323***	-7.573***	-11.700***	-3.373***	-18.870***	-1.106***
	(0.409)	(0.344)	(7.448)	(0.649)	(0.442)	(0.578)	(0.384)	(2.382)	(0.091)
β_N	-5.302***	-4.581***	53.065***	-19.995***	-53.295***	-11.227***	-22.494***	-25.708***	-0.784***
	(0.726)	(0.771)	(11.860)	(2.541)	(2.635)	(1.793)	(2.198)	(7.662)	(0.214)
ln*Dist.*	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Border.Reg	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***	1.951***
	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)
Border.Country	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***	2.693***
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
Adj.Region	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***	0.682***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Adj.Country	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***	0.231***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Observations	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065	1,267,065

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * p<0.10, ** p<0.05,*** p<0.01

Table 6. International and intranational elasticities of trade (pool sectors. Country to region).

Sector Variable	A01	A02	B	C10-C12	C13-C15	C17	C19	C20	C21	C22
$\sigma_F = 1 - \beta_i$	6.463***	5.329***	4.913***	17.681***	30.681***	9.674***	11.379***	4.078**	–	42.905***
$\sigma_N = 1 - (\beta_i + \beta_N)$	12.101***	7.243***	6.315***	31.065***	5.790***	20.717***	–	0.639**	–	9.163**
$\sigma_F \overset{\geq}{<} \sigma_N$	<	<	<	<	>	<	=	>	–	>
β_i	-5.463***	-4.329***	-3.913***	-16.681***	-29.681***	-8.674***	-10.379***	-3.078**	-1.351	-41.905***
	(1.899)	(1.379)	(1.188)	(3.554)	(7.772)	(1.742)	(2.154)	(1.497)	(0.824)	(13.156)
β_N	-5.638***	-1.914***	-1.402***	-13.384***	24.891***	-11.043***	2.664	3.439**	0.804	33.742**
	(2.023)	(0.722)	(0.437)	(3.308)	(6.906)	(3.373)	(2.190)	(1.465)	(0.649)	(16.031)
ln*Dist.*	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Border.Reg	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***
	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)
Border.Country	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***
	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)
Adj.Region	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***
	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)
Adj.Country	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**
	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)
Border.Reg.Count.	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***
	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Observations	235,068	235,068	235,068	235,068	235,068	235,068	235,068	235,068	235,068	235,068

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * p<0.10, **p<0.05,*** p<0.01

Sector Variable	C23	C24	C25	C26	C27	C28	C29	C30	C31_C32
$\sigma_F = 1 - \beta_i$	10.115***	11.881***	183.663***	9.972**	12.317**	35.928***	10.186**	39.458***	3.272***
$\sigma_N = 1 - (\beta_i + \beta_N)$	17.942***	6.246**	–		42.879***	20.925*	–	–	1.876***
$\sigma_F \overset{\geq}{<} \sigma_N$	<	>	=	>	<	>	=	=	>
β_i	-9.115***	-10.881***	-182.663***	-8.972**	-11.317**	-34.928***	-9.186**	-	-2.272***
	(2.019)	(2.899)	(29.020)	(4.254)	(4.907)	(10.334)	(3.778)	(9.869)	(0.528)
β_N	-7.827***	5.635**	28.998	0.968	-30.562***	15.003*	-1.744	5.841	1.396***
	(6.026)	(2.309)	(74.491)	(4.640)	(13.298)	(4.671)	(3.320)	(14.313)	(0.663)
ln*Dist.*	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***	-0.316***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Border.Reg	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***	6.551***
	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)	(0.481)
Border.Country	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***	5.324***
	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)	(0.544)
Adj.Region	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***	0.505***
	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)	(0.128)
Adj.Country	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**	-0.371**
	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)
Border.Reg.Count.	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***	-0.366***
	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Observations	235,068	235,068	235,068	235,068	235,068	235,068	235,068	235,068	235,068

Notes: Imported and imported fixed effects; Errors clustered by region-pair; Significance Levels: * p<0.10, **p<0.05,*** p<0.01

6. Conclusions

This study introduces the theory and practice allowing the estimation of two intertwined measures of import elasticity of substitution. Within the existing regional general computable general equilibrium models (RGCE), normally covering administrative units belonging to a common area characterized by a single market (e.g., the EU), it is possible to differentiate demand equations for domestically produced goods (i.e., within the same region), those imported from regions within the same country (national imports), and, finally, those sourced from regions situated in other countries (foreign imports). This gives rise to home-national (or home-regional) and home-foreign elasticities of trade, whose value has never been jointly determined. Knowledge of both elasticities is critical for the correct calibration of RCGE models and subsequent policy analyses. To the extent that RCGE models disregard the reality behind these two levels of trade flows by adopting single-valued elasticities of trade, their results will be biased, thereby compromising the recommendations for trade policy. In particular, the welfare effects of trade (and transport) policies critically depend on their values. However, within RCGE modeling, there is no need to adopt single-valued elasticities drawn from international trade literature, since it is possible to define and calculate both levels of elasticities.

We develop a three-tier theoretical model based on the CES utility function specification that provides the microeconomic foundation for the gravity equations from which these national (or regional) and foreign (or international) elasticities of substitution can be recovered. The equations are then econometrically estimated through poisson pseudo maximum likelihood (PPML) methods using EU trade data. The theoretical model is consistent with the analytical framework of the RHOMOLO curated by Joint Research Center of the European Commission, while the datasets for the key and ancillary variables are obtained from the model's databases. The reason is that full compatibility is required if this model is to benefit from our research by straightforwardly adopting the estimated values of trade elasticities in the necessary calibrations. However, we contend that these estimated elasticities can be useful to all sort of RCGE models which routinely adopt values corresponding to international studies and that cannot differentiate between the two levels of import substitutability. A crucial issue regarding the data is the construction of a reliable transport cost measure, since this is the key variables from which the trade elasticities are recovered (as opposed to international trade models where tariffs serve to this identification purpose). We calculate a very detailed matrix of generalized transport costs that accounts for the actual road infrastructure, optimal vehicle size depending on shipping distance and urban layout, as well as type of cargo. Coupled with information on average loads and unit price, this allows us to calculate specific origin-destination iceberg transportation costs. Such detailed methodology for calculating transportation costs has never been brought into the international trade literature related to the estimation of trade elasticities.

We explore alternative estimation strategies based on the standard sector by sector (individual) estimation of both sets of elasticities as well as pooling the data by sectors so as to take advantage of larger sample sizes. The results from the individual estimations are unsatisfactory as the values for both sets of elasticities cannot be recovered. Indeed we cannot reject the hypothesis that national and foreign trade elasticities are equal. This limitation is overcome when adopting the pooled regression, as all elasticities exhibit the correct sign, sensible magnitudes and, more importantly, are statistically significant. In our preferred specification the average value for the home-national trade elasticities is 26.9, ranging from 1.8 to as much as 134.0, while average home-foreign trade elasticities is 17.9, ranging from 2.1 to 187.2. Consequently, we conclude that national trade elasticities are in general larger in magnitude that their foreign counterparts. To check the robustness of the results we also perform the same set of regressions but aggregating at the country level the imports from regions belonging to foreign countries. Data aggregation has significant effects on the national elasticities of trade. While foreign elasticities remain basically the same, the latter exhibit much lower values, to the extent that the turn our smaller than their foreign counterpart. Thus, the level of data aggregation is not neutral.

We conclude encouraging researchers involved in regional GCE modeling with different spatial levels of trade flows and transportation costs to study the present proposal and explore the need to differentiate between national and foreign trade elasticities. To the extent that their magnitudes and differences between the two are statistically significant, an important feature of the trade flows between regions would be missing. Adopting theoretical frameworks where this reality is overlooked could seriously compromise the reliability of the models and our confidence in the policy recommendations derived from them. To address this void we provide the modeling tools and econometric methods that enable the implementation of the current proposal, which should prove useful in the theory and practice of regional GCE modeling by improving the characterization of consumers' behavior, while emphasizing the need for a careful calculation of transport costs. Future research will address how the improvements we propose alter the results obtained from a given policy experiment within the current and our newly proposed analytical framework. One key experiment would be determination of the effect of transport infrastructure investments reducing transport costs on social welfare.

Acknowledgements

We are very grateful to Giovanni Mandras for providing us with the trade flows data used in this research. The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Bibliography

Aspalter, L., (2016) Estimating Industry-level Armington Elasticities for EMU Countries. Department of Economics Working Paper No. 217. WU Vienna University of Economics and Business, Vienna.

Bilgic, A., King, S., Lusby, A and Schreiner, D.F. (2002) "Estimates of U.S. Regional Commodity Trade Elasticities of Substitution," *The Journal of Regional Analysis and Policy*, 32(2), 79-98.

Bröcker, J. (2015) "Spatial computable general equilibrium analysis," In: Karlsson, C., Andersson, M. and Norman, T. (eds.), *Handbook of Research Methods and Applications in Economic Geography*. Cheltenham: Edward Elgar.

Burstein, A., and Melitz, M. (2013) "Trade Liberalization and Firm Dynamics". In: Acemoglu, D., Arellano, M. and Dekel, E. (eds.), *Advances in Economics and Econometrics: Tenth World Congress* (Econometric Society Monographs, pp. 283-328). Cambridge: Cambridge University Press.

Bühlmann, P. and van de Geer, S. (2011). *Statistics for high-dimensional data: Methods, theory and applications*. Berlin: Springer-Verlag.

Di Comité, F., Kancs, d'A. and Lecca, P. (2017) "Modeling agglomeration and dispersion in space: The role of labor migration, capital mobility and vertical linkages," *Review of International Economics*, 26: 555–577.

Dijkstra, E.W. (1959) "A note on two problems in connexion with graphs," *Numerische Mathematik* 1(1), 269-271.

Dix-Carneiro, Rafael (2014) "Trade Liberalization and Labor Market Dynamics," Econometrica, 82 (3), 825-885.

Durà A., Camonita F., Berzi M. and Noferini A. (2018). Euroregions, Excellence and Innovation across EU borders. A Catalogue of Good Practices. Barcelona, Department of Geography, UAB, Catalonia, Spain.

Feenstra, R. (2016) *Advanced International Trade. Theory and Evidence*. 2nd Ed. Princeton: Princeton University Press.

Feenstra, R.C., Luck P.L., Obstfeld M., and Russ K.N., (2014) In Search of the Armington Elasticity. National Bureau of Economic Research (NBER) Working Paper 20063.

Fieler, A.C. (2011) "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 79(4), 1069-1101.

Francois, J., and Martin, W. (2013) "Computational General Equilibrium Modelling of International Trade". In: Bernhofen D., Falvey R., Greenaway D., Kreickemeier U. (eds) *Palgrave Handbook of International Trade*. Palgrave Macmillan, London

Giuliano, G., Dablanc, L. and Rodrigue, J.-P. (2019) “Freight and the City,” in Conway et al. (edts.), *City Logistics: Concepts, Policy and Practice*, New York: Routledge. https://globalcitylogistics.org/.

Head, K., and Mayer, T. (2009) Illusory border effects: distance mismeasurement inflates estimates of home bias in trade. In: Van Bergeijk, P. and Brakman, S. (eds.) *The Gravity Model in International Trade: Advances and Applications*. Editors. Cambridge: Cambridge University Press.

Head, K. and Mayer, T. (2013) Gravity equations: workhorse, toolkit and cookbook, CEPR Discussion Papers No. 9322, Centre for Economic Policy Research, London, England.

Hillberry, R., and Hummels, D. (2013) “Trade Elasticity Parameters for a Computable General Equilibrium Model.” In: Dixon, P., and Jorgenson, D. (eds) *Handbook of Computable General Equilibrium Modeling, vol. 1*. Oxford. Elsevier B.V.,

Hummels, D. (2001).Toward a Geography of Trade Costs. Purdue University, Purdue, N.

Imbs, J. and Méjean, I., (2010) Trade Elasticity. A Final Report for the European Commission. Economic Papers 432. Directorate-General for Economic and Financial Affairs Publications. Brussels.

Imbs, J. and Méjean, I., (2015) “Elasticity Optimism,” *American Economic Journal: Macroeconomics*, Vol. 7(3).

Irarrazabal, A., Moxnes, A., and Opromolla, L.D. (2015) “The tip of the iceberg: a quantitative framework for estimating trade costs,” *Review of Economics and Statistics* 97(4): 777–792.

Janic, M. (2007), “Modelling the full costs of an intermodal and road freight transport network”, *Transportation Research Part D: Transport and Environment* 12, 33–44.

Lecca, P., Barbero, J., Christensen, M.A., Conte, A., Di Comite, F., Diaz-Lanchas, J., Diukanova, O., Mandras, G., Persyn, D. and Sakkas S. (2018) RHOMOLO V3: A Spatial Modelling Framework, EUR 29229 EN, Publications Office of the European Union, Luxembourg.

Németh, G., Szabõ, L., and Ciscar, J. C., (2011) “Estimation of Armington elasticities in a CGE economy-energy-environment model for Europe,” *Economic Modelling* 28 (4), 1993-1999.

Meyer, T. and Zignano, S. (2011), Notes on CEPII’s Distances Measures: The GeoDist Database, CEPII Working Paper No. 2011-25, Paris.

Olekseyuka, Z. and Schürenberg-Froschb, H., (2016), “Are Armington elasticities different across countries and sectors? A European study,” *Economic Modelling,* 55(2), 328-342.

Perkmann, M. (2002), “Euroregions: Institutional Entrepreneurship in the European Union,” In M. Perkmann and N.-L. Sum (Eds.), *Globalization, Regionalization and Cross-Border*

Regions (pp. 103–124). London: Palgrave Macmillan UK.

Persyn, D., Díaz-Lanchas, J., and Barbero, J. (2019). Estimating road transport costs between EU regions. JRC Working Papers on Territorial Modelling and Analysis No. 04/2019, European Commission, Seville, 2019, JRC114409.

Romalis, J. (2007), "NAFTA's and CUSFTA's Impact on International Trade," *Review of Economic and Statistics* 89, 416-435.

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

ISSN 0008-

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

Santos Silva, J. M. C., & Tenreyro, S. (2010). "On the existence of the maximum likelihood estimates in Poisson regression," *Economics Letters* 107(2), 310–312.

Santos Silva, J.M.C. and Tenreyro, S. (2015) PPML: Stata module to perform Poisson pseudo-maximum likelihood estimation, Department of Economics, Boston College, Statistical Software Components series, #S458102.

Thissen M., Husby, T., Ivanova, O. and Mandras G. (2019). European NUTS 2 regions: construction of interregional trade-linked Supply and Use tables with consistent Transport flows. JRC Working Papers on Territorial Modelling and Analysis No 01/2019.

Welsch, H., (2008), "Armington elasticities for energy policy modeling: evidence from four European countries," *Energy Economics* 30 (5), 2252-2264.

Zofío, J. L., Condeço-Melhorado, A. M., Maroto-Sánchez, A., and Gutiérrez, J. (2014), "Generalized transport costs and index numbers: A geographical analysis of economic and infrastructure fundamentals," *Transportation Research Part A: Policy and Practice* 67, 141-157

Appendixes

Appendix 1. Correspondence table between the CPA 2.1 and NST 2007 classifications

CPA 2.1	NST 2007	Description CPA 2.1
CPA_A01	01	Products of agriculture hunting and related services
CPA_A02	01	Products of forestry logging and related services
CPA_A03	01	Fish and other fishing products, aquacult. Products, support services to fish.
CPA_B	02-03	Mining and quarrying
CPA_C10-C12	04	Food products beverages and tobacco products
CPA_C13-C15	05	Textiles wearing apparel and leather products
CPA_C16	06	Wood and of products of wood and cork except furniture articles of straw …
CPA_C17	06	Paper and paper products
CPA_C18	06	Printing and recording services
CPA_C19	07	Coke and refined petroleum products
CPA_C20	08	Chemicals and chemical products
CPA_C21	08	Basic pharmaceutical products and pharmaceutical preparations
CPA_C22	08	Rubber and plastics products
CPA_C23	09	Other non-metallic mineral products
CPA_C24	10	Basic metals
CPA_C25	10	Fabricated metal products except machinery and equipment
CPA_C26	11	Computer electronic and optical products
CPA_C27	11	Electrical equipment
CPA_C28	11	Machinery and equipment nec
CPA_C29	12	Motor vehicles trailers and semi-trailers
CPA_C30	12	Other transport equipment
CPA_C31_C32	13	Furniture other manufactured goods

Sectors CPA_A03 and CPA_C16-C18 are not included in the analysis due to the lack of data regarding either transportation and/or trade data.

Note: The correspondence tables between the two classifications can be at the Eurostat's RAMON site: https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL&StrLanguageCode=EN&IntCurrentPage=11

Appendix 2. Economic cost factors for selected vehicles depending on size.

Economic costs	Vehicle HDV (5 axles) Large	 Rigid (3 axles) Medium	 Small (2 axles) Small
Variable costs			
Distance			
Fuel	1.000	0.611	0.317
Tire	1.000	0.911	1.041
Maintenance	1.000	1.282	1.121
Time			
Labor	1.000	1.000	1.000
Amort&Fin	1.000	0.626	0.296
Insurance	1.000	0.715	0.623
Indirect	1.000	0.571	0.445
Fixed costs			
Handling	1.000	0.752	0.309
Tax	1.000	0.910	0.849
Vignette	1.000	0.600	0.600

Note: The standard Heavy Duty Vehicle (HDV) costs reported in Persyn et al. (2019), corresponding to a 40 ton articulated truck, are the baseline for the remaining vehicles.

Appendix 3. Economic cost factors for selected vehicles depending on cargo.

Costs	Standard HDV	Liquid/Solid Bulk Food	Liquid/Solid Tanker Dangerous	Tanker Gas	Liquid/Solid Bulk	Carrier vehicles	Container (Skeletal)
Variable	1.000	1.313	1.270	2.226	1.345	1.222	0.997
Tax	1.073	1.173	1.163	1.058	1.142	0.893	1.073
Vignette	1.087	1.150	1.797	1.000	1.190	1.000	1.087
Handling	1.000	1.000	1.000	1.000	0.000	1.000	1.000