I.S.S.N.: 1885-6888

ECONOMIC ANALYSIS
WORKING PAPER SERIES

Trade margins, transport cost thresholds and market areas: Municipal freight flows and urban hierarchy

♦

Jorge Díaz-Lanchas, Carlos Llano-Verduras and José L. Zofío

Working Paper 10/2013

UNIVERSIDAD AUTONOMA DE MADRID

DEPARTAMENTO DE ANÁLISIS ECONÓMICO: TEORÍA ECONÓMICA E HISTORIA ECONÓMICA

Trade margins, transport cost thresholds and market areas: Municipal freight flows and urban hierarchy

Jorge Díaz-Lanchas[a,b,*], Carlos Llano-Verduras[a,b], Jose Luis Zofío[a]

[a] *Department of Economics. Universidad Autónoma de Madrid, 28049, Cantoblanco, Madrid. Tel.: (34) 91-4976768.*
[b] *Lawrence. R. Klein Institute, Universidad Autónoma de Madrid.*

Abstract

Recent research has determined the existence of a border effect on trade flows within a country associated to agglomeration economies, the size of the spatial unit of reference, as well as to alternative measures of transport costs. Using a micro-database on road freight shipments within Spain for the period 2003-2007, we consistently decompose the total value of municipal freight flows into the extensive and intensive margins at the European Nuts-5 (municipal), 3 (provincial) and 2 (regional) levels and study the impeding effect of actual generalized transport costs (as opposed to proxies given by the standard measures of distance and travel time). Establishing the superiority of this generalized measure of transport costs, we confirm the accumulation of trade flows up to a transport cost value of 330 euros, and conclude that this high density is not explained by the existence of administrative limits (border effects) but to significant changes in the trade flows-transport costs relationship. While this high density of trade coincides with low level administrative borders (municipal and provincial) as there is a positive and significant effect associated to them on all trade decomposition, it is not significant, or even negative, at a larger regional level. To support this hypothesis, we identify significant thresholds in the trade flows-transport costs relationship that are calculated by way of the Chow test of structural change. These breakpoints allow us to split the sample and control for successive administrative borders in both the extensive and intensive margins. Relying on these thresholds we define relevant market areas corresponding to specific transport costs values that portrait a consistent urban hierarchy system of the largest Spanish cities within a radius of about 330 euros, thereby providing clear evidence of the predictions made by the central place theory.

Keywords: Municipal Freight Flows, Transport Costs, Breakpoints, Market Areas, Urban Hierarchy, Central Place Theory.

JEL Classification Numbers: F14, F15, O18

*Corresponding author.

Email addresses: jorge.diaz@uam.es (Jorge Díaz-Lanchas), carlos.llano@uam.es (Carlos Llano-Verduras), jose.zofio@uam.es (Jose Luis Zofío)

1. Introduction

Why and to what extent trade flows tend to agglomerate in some specific places within a country are usually topics left out of the analysis due to the non-existence of very detailed micro-data on interregional trade flows.[1] There are some studies in which trade flows agglomerate as a result either of intermediate inputs flows (Hillberry (2002a); Hillberry and Hummels (2003, 2008)) or because of the location of firms close to a real border as to reduce transport costs when trading with other countries—Hillberry (2002a); Chen (2004)), but there is not a systematic analysis on why trade flows agglomerate in cities within a country, specially using very detailed trade flows. In this sense, the study by Hillberry and Hummels (2008) sheds light on the appearance of an internal "home bias" at the municipal level once we use accurate interregional trade flows and very precise measures of internal transport costs, although they concluded that this "internal home bias" is a *reductio ad absurdum* of the home bias effect.

In this paper we extend the analysis by these last authors incorporating some econometric novelties related to trade and urban economics explaining why trade flows agglomerate in some specific places resulting in a hierarchical pattern of cities. In the process we bear in mind that most of the trade literature focuses on international trade flows, even if the largest share of the trade activity is performed within a country, and specifically, between cities, showing how pertinent the present study is.

Going beyond the study by Hillberry and Hummels (2008), we argue that this "internal home bias effect" is an *illusion* created by the existence of transport costs thresholds which shape a series of trade or market areas driven by the biggest trading-cities within a country. Indeed, in a further step, we present empirical evidence by which this agglomeration of trade flows around bigger cities responds to the existence of a hierarchical urban system as it is predicted by the Central Place Theory—McCann (2001); Parr (2002); Tabuchi and Thisse (2011); Mulligan et al. (2012).

To perform this analysis, we make use of micro-data at the highest possible level of spatial disaggregation corresponding to individual shipments at the Nuts-5 municipal level, for the period 2003-2007. This particular database allows us to determine the existence of a concentration of trade flows in small areas (in this case cities) defining *natural* trade areas in the spirit of the central place and location theory—Tabuchi and Thisse (2011), and whose geographical reach is directly related to this transport costs thresholds, which may even exceed alternative levels of administrative aggregation of spatial units; particularly, Nuts-5 and Nuts-3 territorial units. This allows us maintain that it is precisely the existence of these transport related breakpoints (thresholds)—not always coinciding with relevant administrative geographical limits—what really explains the appearance of an "internal home bias effect".

[1] Recently, there have appeared studies such as Llano et al. (2010); Yi (2010); Llano-Verduras et al. (2011); Yi (2010); Borraz et al. (2012); Yilmazkuday (2012) from an empirical perspective; meanwhile from a theoretical approach Behrens et al. (2013) provide a promising attempt to model how trade frictions affect goods shipped between cities.

This paper presents several empirical novelties that link and extend different literatures: 1) the study of the internal home bias effect (internal border effect) within a country: Hillberry (2002a); Hillberry and Hummels (2003, 2008) against which we argue that administrative borders do not play an effective role in halting trade as impediments have been removed since long as a result of "single-market" agreements; 2) the trading-cities theoretical literature whose empirical contrast is still pending: Anas and Xiong (2003); Behrens (2005); Cavailhes et al. (2007), and to which we provide relevant illustrations and complementary insights; 3) the study of market areas as regional boundaries depending on the geographical reach of trade: Löffler (1998), and, finally, 4), the re-emergence of the Central Place Theory literature—McCann (2001); Parr (2002); Tabuchi and Thisse (2011); Mulligan et al. (2012).

To accomplish these consecutive goals we rely on: 1) the compilation of a very detailed panel dataset (2003-2007) on individual freight trade flows; 2) the definition and calculation of a precise and realistic economic measure of transport costs associated to these flows: the Generalized Transport Cost, GTC, performing better that its usual distance and time counterparts in panel data contexts; 3) the adoption of a new methodological approach within the trade literature based on structural econometric tests (endogenous Chow Test) to determine transport cost thresholds, which in turn 4) allows us to define natural market trade areas at different spatial levels of aggregation for the overall value of trade, and its extensive and intensive margins; 5) the interpretation of results in terms of the central place and location theory and their discussion by way of suitable and novel Geographic Information Systems (Arc/GIS) illustrations; and, finally, 6) the study of all these issues using panel data econometrics that enable us to capture the dynamics of the relationship between trade flows, transport costs and internal border (home bias) effects.

To perform this study we adopt a sequential strategy. The first step requires the compilation of a novel database on road freight shipments consisting of micro-data for individual deliveries between Spanish municipalities.[2] The next step is based on RFTS data reporting the origin and destination of a shipment. For each bilateral freight service we calculate the actual monetary measure of transport costs (i.e., generalized transport cost, GTC) to better capture the dependence of trade flows and its extensive and intensive margins on their transport cost. The GTC corresponds to the minimum cost of joining any origin and destination, defined as the sum of the cost related to distance (e.g., fuel, toll, tires) and time (e.g., salaries, insurance, taxes). To make these calculations we have resorted to programming techniques using Geographic Information Systems (Arc/GIS) that allow the optimization of the least cost routes through the existing road network in the years 2003-2007. The advantage of the GTC over its distance and time proxies is that it accounts for their associated transportation operating costs—see Combes et al. (2005); Zofío et al. (2011), thereby capturing the change in the transport service market as a result of changes in input prices (e.g., particularly fuel)

[2] The *Road Freight Transportation Survey*, RFTS, differs in several ways from the American *Commodity Flow Survey* undertaken through a partnership between the Census Bureau and the Bureau of Transportation Statistics (DOT). Relevant for this study is that in the latter case, the surveyed statistical units are production establishments (wholesalers and retailers), while the RTFS surveys road freight companies producing the transportation service.

and regulatory conditions (e.g., related to wages in the labor market). Therefore, in contrast to all previous studies using the standard and non-monetary transport cost proxies of distance and time, we introduce a *real* Euro measure of the spatial related frictions affecting trade.

Subsequently, to study the dependence of trade flows on transport costs and the magnitude (either real or illusory) and significance of the "internal home bias effect" we rely on the pseudo-poisson maximun likelihood (PPML) estimator proposed by Santos Silva and Tenreyro (2006, 2010, 2011), as the most efficient way to control for pervasive zero value trade flows and heteroskedasticity problems within the gravity model. As anticipated, the statistical information reported in the *RFTS* allows us to decompose the value of trade flows into the extensive and the intensive margins following the trade flows definition proposed by Hillberry and Hummels (2008). The results obtained using the PPML estimation including time-varying origin and destination fixed effects, show that "municipal boundaries" have a stronger impact on trade flows and trade margins than the results reported by Hillberry and Hummels (2008); i.e., trade inside municipalities or between contiguous municipalities is much more important than trade between non-contiguous or long distant municipalities, especially in the extensive margin. But this is explained by the changing pattern (elasticity) of the effect of transport costs on trade, which is particularly intense for short distances (Euros)—i.e., trade sharply decreases within the administrative boundaries of a municipality—rather than any border effect impediment. In this sense, considering administrative boundaries not as borders haltering trade, but as likely distances where the trade flow-transport cost relationship changes, they do reflect different effects on trade. Indeed the borders between provinces (Nuts-3) have a seizable effect on all the trade decomposition variables, but lower than the municipal level (Nuts-5), while regional borders (Nuts-2) have no significant or even an inverse effect on trade flows.

These findings corroborate the idea that the impact of high administrative boundaries (Nuts-2) on trade flows is not as important as the trade literature has emphasized: Wei (1996); Head and Mayer (2002); Anderson and van Wincoop (2003); Evans (2003); Brülhart and Trionfetti (2009); Benedictis and Taglioni (2011). Particulary, if we decompose the cumulated impact comprised within high administrative boundaries (i.e., national units) into lower regional borders, we obtain a concentration of trade flows within boundaries which is higher than expected; but almost completely determined by the border corresponding to very reduced administrative boundaries (e.g., within the municipal level), with the rest of the effects turning out to be marginal in relative terms. Indeed, this result suggests that the trade-agglomeration effect at the municipal level is more related to agglomeration economies (Hillberry (2002a); Chen (2004); Puga (2010)) than with differences in consumer preferences for local varieties in detriment to those produced abroad, i.e., idiosyncratic demand, cultural differences, and trade inertias (Wei (1996); Evans (2003); Brülhart and Trionfetti (2009). Furthermore, we argue that these agglomeration economies only take place around bigger cities because of the existence of a urban hierarchy system that emanates from the Central Place Theory (Parr (2002); Tabuchi and Thisse (2011); Mulligan et al. (2012)). Hence, the so called "home bias effect" arises at the municipal level because higher-order cities act as supply center either for its metropolitan area or for the lower rank surrounding cities. A relationship not explored previously in the trade and urban literature.

To study this hypothesis we determine to what extend the extensive margin (number of shipments) are geographically located at very short distances, while the intensive margin (average value per shipment) remain basically constant, reflecting that the effect of transport costs on each component is quite different. To achieve this goal, we introduce a new methodology in the trade literature proposed by Berthelemy and Varoudakis (1996) for endogenous economic growth models. These authors conjecture the existence of different structural models which appear at specific *breakpoints*. These breakpoints could exist within a cross-section, in contrast to the structural change in time-series models, and are endogenously determined by the variables in the model. To pinpoint this threshold values, they proposed an endogenous Chow Test which divides the cross-section sample in two sub-samples in order to determine if for the transport cost variable there exists a structural breakpoint for the endogenous trade flows in our model. In the current framework, as trade flows are geographically localized in terms of low transport cost (short distances), we perform this test to determine if GTCs are conditioning particular trade pattern. Our hypothesis is based on the idea that transport costs and GDPs in the gravity equations cannot have the same effect on trade as the geographical reach of trade associated to transport cost increases. We conjecture that this localized trade pattern in short distances is due to the extensive margin, but once trade flows achieve a relevant breakpoint (a GTC threshold), the extensive margin and the intensive margin present two diverging and even opposite patterns; i.e., the intensive margin begins to gain relevance while the extensive margin becomes flat at rather low values. That is why we cannot estimate the effect of administrative boundaries over trade flows as a *mean effect* for the overall sample because not all the administrative boundaries appear in every shipment, neither the extensive and intensive margins always dominate at the same GTCs ranges, and are not equally important.[3]

Furthermore, we have performed the structural Chow Test several times in order to determine what are the transport cost thresholds that condition interregional trade flows, and split our PPML regressions in terms of these new thresholds—in contrast, for example, to Eaton and Kortum (2002), who divide trades flows considering arbitrary distances. With these new breakpoints, we conclude that the home bias (border) effect is not unique, and it does not have an average, nor a linear impact on trade as the trade literature emphasizes, as it spills out over consecutive administrative boundaries because these breakpoints are not affected by arbitrary administrative borders. Moreover, our findings present some promising results as the trade costs estimation is very penalizing on short distances, while it becomes less detrimental for long distances or even not significant for medium distances. Indeed, our gravity-threshold model losses explanatory power for very long distances, indicating the existence of other factors affecting long-distance trade than the traditional ones.

Beyond the determination of these transport cost thresholds we argue that they define *natural trade or market areas* within a country, as they are based on objective statistical

[3]For instance, if a shipment is delivered at a cost of 500€ (about 455 km or 5 hours at the legal speed limit of 90 km/h), we should analyse what margins predominate and what are the spatial limits in which each one of them takes place (Nuts 5, 3 or 2).

criteria; i.e., the application of the endogenous Chow Test obtained directly from our gravity model. Since to the best of our knowledge, this methodology on structural breaks has never been used in trade literature, we provide a relevant illustration of its potential in relation to the Central Place Theory (CPT) models in terms of market areas. Different thresholds create concentric *iso*-CGT rings representing the geographical reach of trade for a location (city), which is increasing on its size and eventually overlaps with the market area of other locations. These market areas naturally define a urban hierarchy providing an empirical justification of the central place theory.

Moreover, according to Parr (2002), O'sullivan (2003) and Mulligan et al. (2012), although central place theory predicts the existence of market areas and explains how cities supply their surrounding areas by way of trade flows, the existing empirical studies fall short from this reference, since they provide illustrations based on population agglomeration rather than actual trade flows. In this work we propose a new mapping representation based on municipal product diversification, and how these calculated market areas shape the urban hierarchy proposed by the CPT by which larger municipal units (*Rank* 1 and *Rank 2* cities) determine an area of influence which is more extensive than the one obtained for smaller cities (*Rank 3* or *Rank 4* cities). The empirical regularities that we find are a promising result for future research linking trade and central place theories.

Once the motivation of the study and the review of the literature have been introduced, we lay out the structure of the article. The next section discusses the database on trade shipments and generalized transport costs and justifies why the GTC is the most suitable measure of transport cost in a time-series/cross-section panel data framework. Section 3 presents the decomposition of the value of trade into the extensive and intensive margins, analyzes their values based on nonparametric kernel density distributions, and shows the results obtained with different specifications of the gravity model using the pseudo-poisson distribution. Here we discuss the determinants of the value of trade and its margins in terms of our monetary measure of transport costs. Section 4 discusses the structural breakpoint methodology based on the endogenous Chow Test that allows determination of transport cost thresholds, and successively replicate the PPML regressions for each range of trade values according to these thresholds. In section 5 we use the results obtained with the gravity model when regressing trade by transport cost thresholds to provide a consistent illustration of the geographical reach of trade and the emergence of natural market areas. We interpret and illustrate our results in term in terms of central place theory, mapping a hierarchical system of cities and their overlapping regional boundaries. The last section draws relevant conclusions.

2. Data

2.1. Trade value data: The road freight transportation survey

For this study we rely on a micro-database on shipments by road within Spain during the period 2003-2007 elaborated within the research project C-intereg.[4] This database is

[4] A complete description of the methodology and relevant data are presented in www.c-intereg.es.

based on the annual Road Freight Transportation Survey, RFTS, compiled by the transport division of the Ministerio de Fomento, which randomly surveys a sample of freight companies and independent truckers, with vehicles over 3.5 tons and operating within the national territory. The database includes information about the characteristics of the vehicle and shipments, such as the number of tons carried out by the truck,[5] the number of shipments between the origin i and the destination j, the type of product,[6] the operations performed by the truck in each shipment,[7] as well as the actual travel distance in kilometers between the geographical origin and destination of each shipment (recorded at the Nuts-5 municipal level). As a result, for each shipment, it is not necessary to approximate the distance as done in other studies working with databases that record the origin and destination of the shipment by municipal or ZIP codes (e.g., distance between the centroids of these areas), as the true *door-to-door distance* travelled by the vehicle is reported.[8] Therefore, and thanks to this distinctive feature of the database, we can also research intra-municipal trade flows; a relevant micro level flow that is normally left out of the analysis when the database does not record the real distance of shipments.

For the 2003-2007 period, the database contains more than 1.890.000 records involving, on average for all these years, 7.178 municipalities of origin from where a freight service is made and 7.913 destination municipalities. However, most of these origins and destinations are municipalities with little relevance in terms of population and trade volumes, so estimation results would be biased. Therefore, we have filtered the sample considering only municipalities that, on average for the period, had over 10.000 inhabitants. As a result, we get a sample of 633 municipalities whose trade volume by road represents a 75.5% of the total.[9]

Since the survey does not provide information about the value of the traded goods, product prices (in euros per ton) are needed so as to obtain a magnitude of the total value moved once they are multiplied by the tons carried. These prices are not available at the municipal level (Nuts-5) in any of the official databases because of statistical confidentiality. To overcome this limitation, we rely on an alternative database that contains bilateral trade flows between Spanish regions (*Comunidades Autónomas*) (Nuts-2) and provinces (Nuts-3), compiled by the C-intereg project.[10] With this database we calculate a price vector,

[5]This corresponds to the real load of the truck in tons. Note that the truck load may range from zero to 100%, so the database may record empty truck movements as a result of the vehicle moving to a destination where it will be eventually loaded.

[6]In the micro-database, commodities are classified attending to the Eurostat classification NST-R which differentiates between 180 products.

[7]"Operations" refers to loads or downloads carried out by the truck in each shipment.

[8]Unfortunately, the database does not compile any information on the firms involved in the shipments. However, since we know the type of product being shipped we can approximate the production sector of the firm.

[9]In this sample, we eliminate the trade flows with the Spanish islands and the autonomous cities of Ceuta and Melilla in North Africa, as we study trade flows by road.

[10]This database is supposed to be, as far as we know, the largest database on interregional trade flows estimated in Spain. It includes bilateral trade flows specifying the region (Nuts-2) and the province (Nuts-3) of origin and destination, both in tones and euros. For further information see Llano et al. (2010).

measured in euros per ton, for the whole period. However, these prices are calculated at the provincial level. Therefore, we are forced to assume that prices at the municipal provincial levels are equal. This assumption implies that the pricing rules determining their level at municipal level, i.e., costs and mark-ups over costs, are similar to those observed at provincial level, e.g., similar labor and intermediate costs.

2.2. The generalized transport cost and its distance and time proxies

Another novel aspect of this study is the use of a real (monetary valued) measure of transport cost that clearly improves those normally used as approximations mainly geographical distance. This variable corresponds to a Generalized Transport Cost (GTC) definition corresponding to the least cost itineraries between an origin and a destination. The GTC is calculated using GIS software (Arc/GIS) with the digitalized road network, as discussed in Zofío et al. (2011). GTCs differentiate economic costs related to both distance and time. The distance economic cost (euros per kilometer) includes the following variables: Fuel costs (fuel price); Toll costs (unit cost per km, multiplied by the length of the road); Accommodation and allowance costs; Tire costs; and Vehicle maintenance and repairing operating costs. On the other side, the time economic cost (euros per hours) include the following variables: Labor costs (gross salaries); Financial costs associated to the amortization; Insurance costs; Taxes; Financing of the truck (assuming that it remains operative only for a certain number of hours/year); and indirect costs associated to other operating expenses including administration and commercial costs.[11] Just as in any personal GPS the minimum distance and time routes do not generally coincide, the minimum GTCs route calculated with Arc/GIS differs from their proxies.

In contrast with other studies that use national level operating costs for GTC measurement, the GTCs employed in this study are calculated considering prices at the provincial level; specifically those observed in the province where the shipments originates.

To test the robustness of results and determine how the usual proxies of transport costs perform when confronted to a GTCs measure in gravity equations, we carry out the same set of regressions with the partial variables of real distance (km) and time (hours) that are comprised within the GTC definition once weighted by their corresponding costs and assuming a cost minimizing behavior on the part of firms. Particularly, in the remaining of the paper we consistently report in the main text kernel regressions, econometric estimations, breakpoint thresholds and market areas based on GTCs, while in the appendices we present the results obtained when using its distance and time proxies instead. We highlight that the RFTS database reports the precise door-to-door distance reported by the freight company for each shipment, offering a new level for intra-city or inter-city transport cost (trade within municipalities), which improves in itself the distance measures used in other studies that are based on area centroids. As for the time taken to deliver shipments, it is normally considered as a measure of road network efficiency in terms of speed and congestion, as it captures the prevalence of high capacity roads. The time taken to travel the distance reported in the

[11] The minimum cost itinerary among the set of possible itineraries defines as follows: $GTC_{ij} = min(DistCost_{ij} + TimeCost_{ij})$

	GTC	Distance	Travel Time
GTC	1.00	—	—
Distance	0.91	1.00	—
Travel Time	0.95	0.9	1.0

Table 1: *Correlation between GTC, Distance and Travel Time.*

database is computed taking into consideration the physical attributes of the arcs belonging to the itinerary between the origin and the destination, including distance, road type and road gradient.[12] We also consider relevant legislation such as speed limits, mandatory stops that truck drivers must observe so as comply with European Union safety regulations, etc.

Both the GTC and time measures decrease over the years because of the improvements in road infrastructure as well as changing regulations. On the contrary, distance remains mainly unaffected by road improvements (and it could even increase as a result of business by-pass routes). As shown in the next section, in the Spanish case distance does not vary significantly since it is unaffected by road improvements (e.g., enlarging roads with 2x1 lanes to 2x2 highways). Because of this lack of variability in the distance variable over time, even if it may constitute a good proxy for transportation costs in cross section studies, it is certainly inadequate in panel data studies where transport costs are expected to vary significantly as a result of improvements in road infrastructure and changes in operating costs. Additionally, while the time proxy for transports costs captures the improvements in road infrastructure, it cannot account for changes in the operating costs. This justifies the choice of GTCs as the only suitable transport cost measure in panel data gravity studies as it is the only variable capturing all dimensions affecting their change over time.[13]

The cross-section correlation between the three types of transport costs: Distance, time and GTC, is presented in Table 1. As expected, the three measures of transport costs are highly correlated, but each of them captures different dimensions of transport costs, i.e., geography (distance), road efficiency (travel time) and actual economic costs (GTC).

3. Trade flows decompositions: The extensive and intensive margins

Taking advantage of our detailed micro-dataset on trade, we decompose the value of trade flows into the extensive and the intensive margins. This procedure allows solving potential specification errors in the gravity model when trade flows are not decomposed into these two margins and when analyzing how trade barriers (frictions) affect them (Melitz (2003); Felbermayr and Kohler (2006)). With this purpose in mind, we rely on Hillberry

[12]The road gradient refers to slope. It differentiates between three degrees of slope, flat: 0%-5%; mild moderate: 5%-15%; high: more than 15%.

[13]See Zofío et al. (2011) for a detailed discussion on how the variation of GTCs can be consistently decomposed into infrastructure and cost components using the economic theory approach to index numbers.

and Hummels (2008) and define the total value of shipments between each origin-destination ij pair by T_{ij}, a which can be decomposed in the following way:

$$T_{ij} = N_{ij}\overline{PQ}_{ij} \tag{1}$$

where N_{ij} represents the total number of shipments (extensive margin) and $\overline{PQ}_{ij}$ is the average value per shipment (intensive margin). At a second level, the previous expression can be further broken down so that the total number of shipments N_{ij} equals the number of commodities (k) sent within the same pair ij (N^k_{ij}), multiplied by its frequency or *trading pair* (F); that is, the average number of shipments per commodity per ij (N^F_{ij}):

$$N_{ij} = N^k_{ij}N^F_{ij} \tag{2}$$

With this expression, the extensive margin is decomposed according to the *product extensive margin* (N^k_{ij}) and the *product intensive margin* (N^F_{ij})—Mayer and Ottaviano (2008). Meanwhile, the intensive margin can be decomposed into the average price ($\overline{P}_{ij}$) and the average quantity ($\overline{Q}_{ij}$) for each pair ij:

$$\overline{PQ}_{ij} = \frac{(\sum_{s=1}^{N_{ij}} P^s_{ij}Q^s_{ij})}{N_{ij}} = \frac{(\sum_{s=1}^{N_{ij}} P^s_{ij}Q^s_{ij})}{(\sum_{s=1}^{N_{ij}} Q^s_{ij})}\frac{(\sum_{s=1}^{N_{ij}} Q^s_{ij})}{N_{ij}} = \overline{P}_{ij}\overline{Q}_{ij} \tag{3}$$

Additionally, to obtain regression results based only in physical units (tons) leaving aside prices, we also consider as dependent variable the shipped quantity:

$$Q_{ij} = (\sum_{s=1}^{N_{ij}} Q^s_{ij}) = \frac{\sum_{s=1}^{N_{ij}} \overline{Q}_{ij}}{N_{ij}}N_{ij} \tag{4}$$

In the database the observations are recorded according to each origin-destination pair (ij), type of commodity transported and year.[14] To obtain yearly values of trade for each ij we aggregate observations such that, following (3), we calculate the average quantity ($\overline{Q}_{ij}$) and the average price ($\overline{P}_{ij}$) for all the shipments between each ij, thereby obtaining the average value per shipment by multiplication, ($\overline{PQ}_{ij}$). Afterwards, we multiply the average value per shipment by the total number of shipment (N_{ij}) between ij obtaining the total value of shipments (1). Additionally, we calculate the maximum number of different commodities transported between each ij and multiply it by its frequency (average number of shipments per commodity), so as to obtain (2). For the total trade in quantities (4), we multiply the average quantity moved between ij by the extensive margin (2). As for the GTCs, distance and time values, we remark that average values may change yearly, particularly for intra-municipal shipments that are surveyed on a *door-to-door* basis; i.e., from one facility (establishment) to another within the same municipality, and the sample

[14] Commodities are classified in ten groups, from agricultural products to manufactured goods, including products such as metallurgical, minerals, chemicals and fertilizers, and heavy machinery.

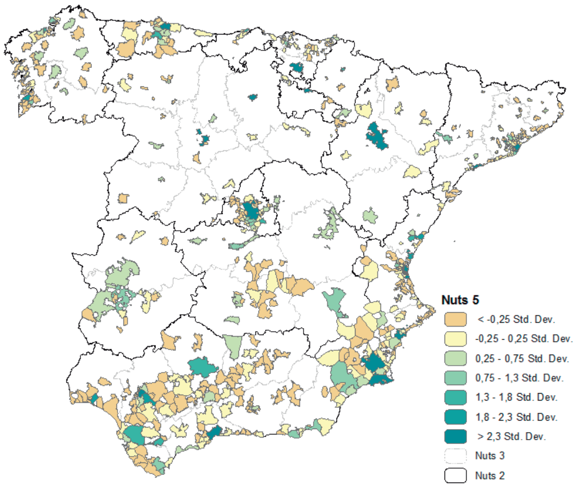


Figure 1: *Standard deviation of the Total Trade for the 633 municipalities. Average Values (Exports+Imports). Period 2003-2007.*

changes every year. Finally, there is a total of 75,897 observations for the whole period after removing empty shipments.[15]

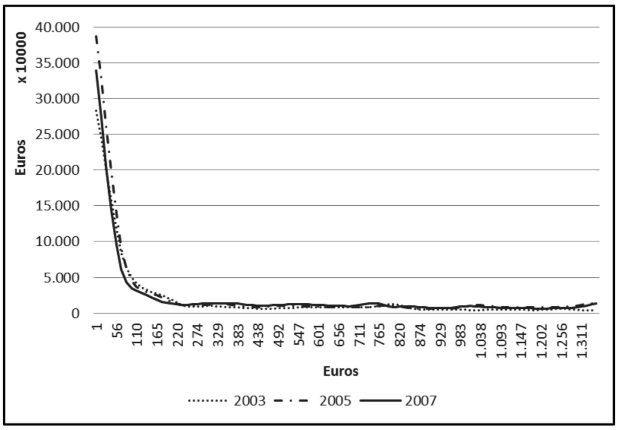
3.1. Descriptive analysis and Kernel regressions

Figure 1 shows the 633 municipalities that have been considered in the final regressions. The map shows the standard deviation of the municipal total trade value, i.e., exports plus imports, both in average values during the whole period. This sub-sample captures all the trade between the largest cities (and metropolitan areas) in Spain. It also includes cities where main ports are located, as areas with high levels of trading activity.

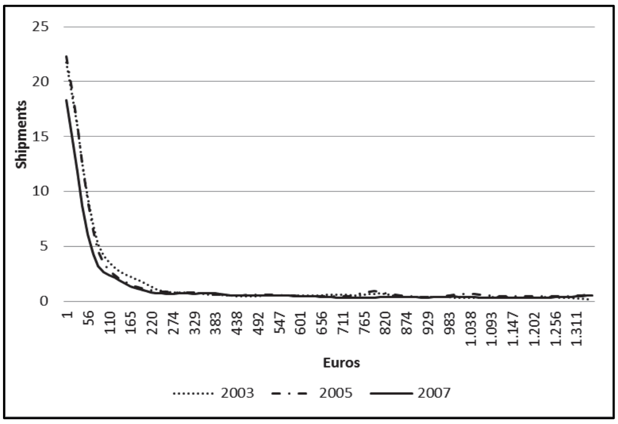
The largest shipments are delivered from the most populated areas with the highest levels of economic activity (Madrid, the Mediterranean, and Basque Country), while less populated areas (south-west and north-west areas) only record trade around the larger cities. Also, trade volumes follow the corridors corresponding to major high capacity roads, indicating the strong inertia between trade and road infrastructures; i.e., firms choose locations with large accessibility defined in terms of market potential Duranton et al. (2013). We use a non-parametric estimation (kernel analysis) to study how each trade variable in the decomposition behaves when considering the alternative measures of transport costs, i.e., the the GTCs (in euros), the actual distance travelled by the truck (in kilometers) and travel time (in minutes), between the 633 municipalities. In the main text we present exclusively the figures considering transport cots (GTCs) in three years intervals: 2003, 2005 and 2007, while in the Appendix B and C we portrait equivalent kernel regressions considering the

[15]At this level of disaggregation, zero trade flows does not represent trade as they correspond to empty truck trips. Similarly, we eliminate zero price observations representing special goods for which no value is reported. These products include: packaging, empty boxes, weapons. Empty flows represent a large percentage of the observations in the sample (around 44%), as a result of unbalanced trade flows within locations, while special goods represent only the 4.5%.

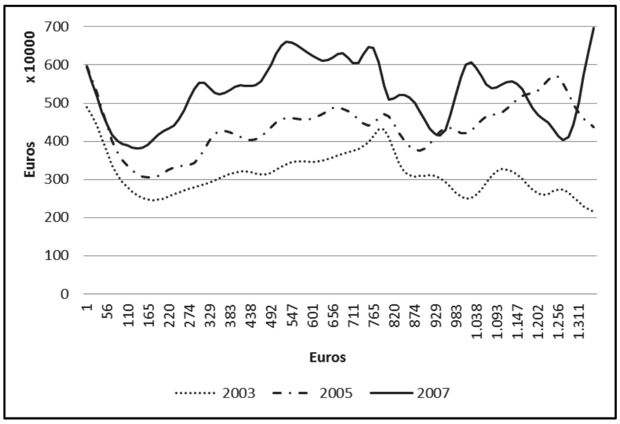
distance and time transport cost proxies. The first level of the trade decomposition is illustrated in Figures 2a, 2b, and 2c presenting total value, number of shipments, and average shipment value, respectively. The total value of shipments is lower in 2003 than in 2007, falling sharply in density as transport costs increase (2a). This same pattern is observed in the extensive margin (2b), where the number of shipments drops rapidly for all years as the GTC reaches the threshold of 150 euros (120 km and 90 minutes, respectively), while the intensive margin (2c) shows a trend that even increases in transport costs. This increasing behavior in the intensive margin is due to its composition. As we will see below, prices naturally increase with GTCs (as its distance and time components), while tons drop in density from 1,300 km onwards, reducing the effect of the high prices at very long distances. Particularly, the GTCs reflect a reduction for middle values because of the discrete travel time variable that is used to calculate them.



(a) *Total Value of Trade*



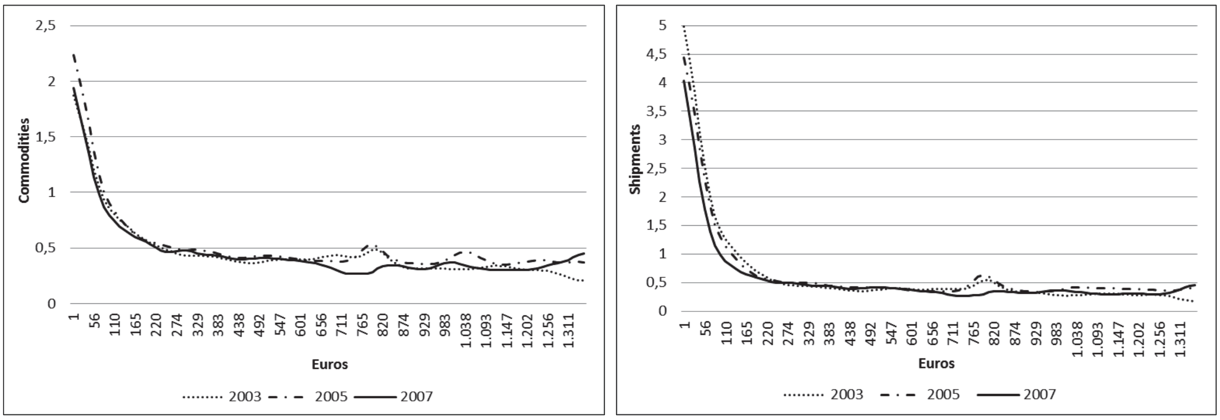
(b) *Extensive Margin*



(c) *Intensive Margin*

Figure 2: *Kernel regressions: Total value of trade, number of shipments (extensive margin) and average value per shipment (intensive margin) on GTC.*

According to the kernel regressions for the total value of trade (Fig.2a) and for the extensive margin (Fig.2b), one can conjecture the existence of a hypothetical threshold or



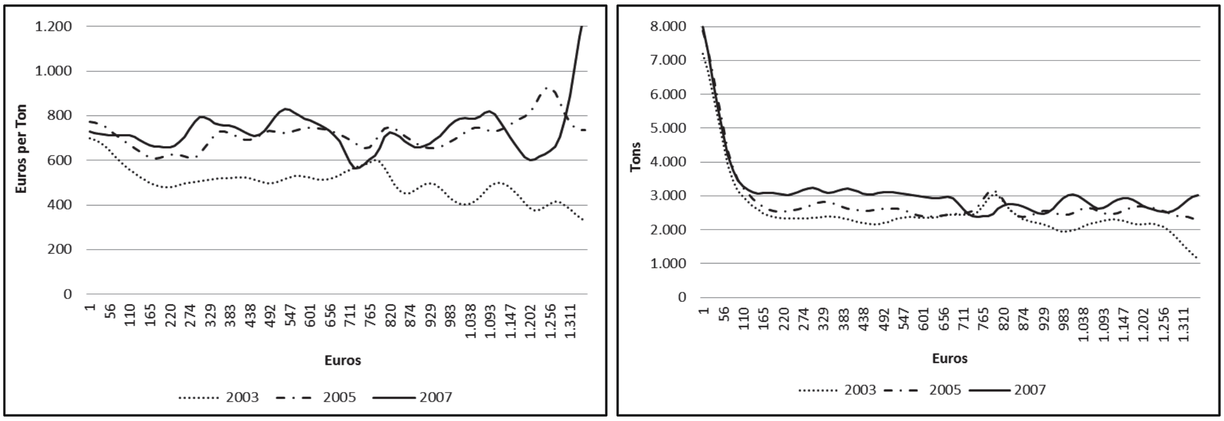
(d) *Number of commodities*

(e) *Number of Shipments per Commodity*

Figure 3: *Kernel regressions: number of commodities and number of shipments per commodity (frequency) on GTC.*

breakpoint threshold where the effect of transport costs on trade flows changes radically. A visual inspection of the figures suggests a breakpoint around 150 euros, which nevertheless is nothing more than an arbitrary value. To rely on objective statistical criteria we perform later a series of Chow structural tests to precisely determine the exact transport cost value (and distance and time proxies) where the model changes structurally. Given this evidence, in the econometric section one should not regress trade flows against transport cost values without considering these two different trade patterns because, otherwise, one would obtain only one *mean effect* of the geographical frictions on trade flows as the trade literature generally presents. That is why if a structural change exists, it is necessary to account for this econometric issue by splitting the sample according to the objective transport costs thresholds that the Chows structural tests yield; thereby obtaining more accurate measures of the effects of trade costs and administrative boundaries on trade flows.

Figures 2 and 3 show the kernel estimations for the second level decomposition -Appendix B and C present equivalent density distribution for the distance and time proxies. Considering the extensive margin decomposition in (2), the number of commodities (Fig. 3d) and its frequency (Fig. 3e) exhibit a remarkably similar pattern and evolution; i.e., they rapidly reach their minimum values for GTC with an increase in density either at middle or high values. Decomposing the intensive margin into its average price and its average quantity, (3), allows us to observe greater price variability, either between years or only in one year. This variability may be due the *specific product* or the *product mix* of the shipment, which results in increasing prices as a result of transport costs increases (Fig. 4f); a sensible result for goods where transport costs make up a large proportion of overall costs, that later on are passed on to prices. Focusing on average tons (Fig. 4g) a relevant fact emerges. Regardless the transport proxy (GTCs, distance and time), all series show a greater density of tons at very short distances, mirroring the extensive margin behavior. Then, they drop sharply to increase again at medium distances. This trend reflects the accumulation of shipments



(f) Average Price (g) Average Tons

Figure 4: *Kernel regressions: average price and average tons on GTC.*

within the main Spanish metropolitan areas (Madrid, Barcelona and Valencia), while beyond these ranges the number of tons reduces until they reach middle transport costs, times and distances from where they increase again. Additionally, it could be indicating that for long distances it is more profitable from a logistics perspective to group shipments and send trucks with a higher capacity and fully loaded, than making many individual shipments with low volume of tons—a behavior reflecting the existence of increasing returns to scale in transportation, McCann (2005).[16] Finally, the kernel regressions for the total trade in physical quantities, (4), not presented to save space, reflect the same evolution and behavior as total trade in monetary units, distance and time, i.e., the total amount falls steeply with increasing transport costs.

3.2. Econometric specification and results

To structure the analysis this section is divided in two parts. In the first one we present the results for the *naïve* gravity model yielding average estimates for the whole sample, and that does not take into account transport cost thresholds when determining the negative impact of transportation, and explore in detail to what extent the effect of administrative boundaries on trade flows have changed over the years. For this purpose we propose a set of regressions using equations (1) through (4) that allows us study how geographical frictions shape trade flows while taking into account different administrative boundaries and measures of transport costs, as well as to test whether these frictions may end up inducing a border effect in each of the different trade margins. In the second part, we propose an analysis based on growth rates to study whether variations in transport costs affect variations in trade flows and the superiority of GTCs over transportation cost proxies.

The set of PPML estimations regressing the value of trade and its components on geographical variables considering as such transport costs, a municipal contiguity variable, the

[16] In the econometric section we study the existence of increasing returns in transport in the intensive (tons) margin.

logarithm of the municipal GDP by origin and by destination, and the three types of administrative boundaries: Regions (Nuts-2), provinces (Nuts-3), and municipalities (Nuts-5), are presented in *Table 2*. While trade and transport costs data has been extensively presented in the second section to motivate their advantage over alternative definitions and proxy variables used in previous studies, we now discuss the remaining variables. To calculate the *Contiguity* variable we use the GEODA software to code if the municipalities shared a common border (*first-order queen contiguity*). It takes the value one if the origin and the destination of the shipment share a border, and also if the shipment takes place within the same municipality to correctly isolate the effect of the municipal boundary (Nuts-5), which captures how important are intra-municipal shipments over inter-municipal ones Hillberry (2002b). Coding shipments within municipalities as adjacent allows the next intradummy (Nuts-5) to isolate the excess local intensity of this type of shipments, relative to the local intensity tied to neighboring municipalities. With respect to municipal GDP, we obtain this variable from the *L. R. Klein Institute* sponsored by the *Servicio de Estudios de La Caixa*.[17] The value index corresponds to the share of economic activity of each municipality over the total national GDP (in per 100,000 terms). This index has a strong correlation with the municipal's market share, so, in order to obtain the municipal GDP, we multiply it by the nominal aggregate national GDP in each year.

For the administrative boundaries, we define three dummy variables as in Requena and Llano (2010). In this sense, the *Nuts_5* variable (municipal boundary) takes the value one if the shipment is performed within the same municipality, and zero otherwise. The *Nuts_3* (provincial boundary) takes the value one if the shipment is carried out between two municipalities which are in the same province but the origin and destination are not the same.[18] Finally, the *Nuts_2* variable (the regional boundary) captures if the shipment takes place between two municipalities which are located in different provinces but they belong to the same *Nuts_2* region; in this case, it will take the value of one and zero otherwise.[19] Additionally, the time dimension is reflected by a dummy variable for each year in the sample Baldwin and Taglioni (2006). Finally, we include time-varying fixed effects by origin and by destination obtaining a regression with more than 5.000 variables—Anderson and van Wincoop (2003); Benedictis and Taglioni (2011). As for the estimation method we rely on the pseudo-poisson maximum likelihood estimation, PPML (see Santos Silva and Tenreyro (2006, 2010, 2011)), considering the endogenous variables in levels. By resorting to the PPML distribution we can accommodate the zero trade flows problem and correct

[17] As one of Europe's leading savings bank and Spain's third largest financial institution, La Caixa sponsors it owns research unit. Particularly, the L. R. Klein Institute elaborates an index based on business (commercial, industrial and services) and professionals taxes collected in each municipality.

[18] If the shipment is within the same municipality, the *Nuts_5* variable will take the value one while the *Nuts_3* and *Nuts_4* variable are assigned a value of zero. Clearly, shipments between municipalities located in different regions are assigned zero values for all administrative boundaries.

[19] As opposed to Hillberry and Hummels (2008) who run separate regressions for the 5 and 3 zip level codes administrative boundaries; our specifications include all boundaries in a single regression. It is understood that they do have a combined and cumulative effect on trade flows, it is necessary to control for each one of them simultaneously, so as to capture interactions and prevent omitted variables biases when regressed separately.

for heterokedasticity.[20] Thus, the final specification to be estimated has the form:

$$\begin{aligned} X_{ijt} = \beta_0 + \beta_1 cost_{ijt} + \beta_2 cost\_sq_{ijt} + \beta_3 contiguity + \beta_4 log\_GDP_{it} + \beta_5\_GDP_{jt} \\ + \beta_6 NUTS\_5 + \beta_7 NUTS\_3 + \beta_8 NUTS\_2 + year + \eta_{it} + \eta_{ijt} + \epsilon_{ijt} \end{aligned} \quad (5)$$

In this specification, the variables *cost* and *cost_sq* denote each transport cost type (GTC, Distance and Time), year corresponds to each dummy year variable in the period, and X_{ijt} stands for each one of the trade decomposition variables already mentioned. The three types of transport costs are considered separately, specifying each cost as a quadratic function to capture the non-linearity between trade flows and transport cost at very short distances as shown by the kernel regressions. All of them are entered in levels to try to capture the non-linearity effect of the distance over trade flows. Thanks to this specification of transport costs, we examine whether there are increasing returns in transport that is; whether a shipment has a negative effect but marginally decreasing with distance. In that case, we would expect a negative sign in the first term of the transport cost variable but a positive sign in the quadratic one—Combes et al. (2005). Finally, we have standardized the three measures of transport costs to avoid problems with the units of measurement to easily compare the results between them.

Table 2 on the next page shows estimates for the first level of trade decomposition variables (the extensive and the intensive margin) taking into account the GTC (distance and travel time regressions are presented in the Appendix A on page 37), plus the additional treatment of trade flows in quantities (tons) to compare it with trade flows in monetary units (total value of trade). Starting with the border or home bias effect while controlling for transport costs, the total value of trade developed within the same municipality (*NUTS_5*) is much greater than inter-municipal trade flows, specifically if we consider GTCs it is 35 times larger.[21] In addition, the higher *NUTS_5* coefficient in the regression of the extensive margin (number of shipments) is indicating that this margin drives intra-municipal trade to a larger extent than the intensive margin (average value per shipment) for all types of transport costs as shown in the kernel regressions. Additionally, if we consider other administrative levels, provincial boundaries (*NUTS_3*) have a much lower effect on trade (4.4 times) than the *NUTS_5* level, while regions (*NUTS_2*) loose importance as administrative boundary, with negligible effects (not even statistically significant for trade in quantities). It turns out then that the border effects for relatively long shipments between regions is nonexistent, supporting the idea that the geographical reach of trade is mainly driven by local markets, i.e., the existence *natural trade areas* in terms of transport costs that we study in the next section. Also, as we are using PPML and time-varying fixed effects by origin and by

[20] Additionally, OLS specifications result in bias estimations Santos Silva and Tenreyro (2006); Martin and Pham (2008).

[21] The PPML regression can be sensitive for extreme observations in the dependent variable causing this high value of the *NUTS_5* level (Van der Marel (2012)). However, as it is a maximum likelihood estimation, is the best technique, in contrast to OLS or quantile regressions, to correctly capture the amount of trade observed in short distance, apart from its econometric advantages (Santos Silva and Tenreyro (2006, 2010, 2011)).

destination, we can correct for the gravity problems evidenced by the literature (Baldwin and Taglioni (2006)); with less biased coefficient estimations in contrast to those obtained by Hillberry and Hummels (2008) and Borraz et al. (2012) using OLS, and indicating a higher impact of the municipal boundary on the trade flows and its margins.

	First Trade Decomposition			**Extensive Margin**		**Intensive Margin**		
VARIABLES	Total Value T_{ij}	Number of Shipments N_{ij}	Average Value per Shipment $\overline{PQ}_{ij}$	Number of commodities N_{ij}^{K}	N. of shipments per commodity N_{ij}^{F}	Price $\bar{P}_{ij}$	Tons $\bar{Q}_{ij}$	Trade in Quantity Q_{ij}
GTC	-0.769***	-0.816***	-0.0788***	-0.508***	-0.531***	-0.157***	-0.282***	-1.039***
	(0.0564)	(0.0287)	(0.0158)	(0.0127)	(0.0155)	(0.0127)	(0.0117)	(0.0460)
GTC Square	0.171***	0.182***	-0.000730	0.110***	0.127***	0.0463***	0.0565***	0.231***
	(0.0161)	(0.00955)	(0.00586)	(0.00485)	(0.00679)	(0.00412)	(0.00454)	(0.0136)
Contiguity	1.325***	1.102***	0.462***	0.637***	0.804***	0.0839***	0.668***	1.230***
	(0.0643)	(0.0337)	(0.0306)	(0.0151)	(0.0264)	(0.0232)	(0.0222)	(0.0485)
log GDP$_i$	0.362***	0.310***	0.215***	0.262***	0.111***	0.178***	0.132***	0.228***
	(0.0235)	(0.0145)	(0.00860)	(0.00532)	(0.00787)	(0.00776)	(0.00589)	(0.0199)
log GDP$_j$	0.372***	0.347***	0.225***	0.262***	0.166***	0.203***	0.137***	0.261***
	(0.0245)	(0.0139)	(0.00824)	(0.00523)	(0.00777)	(0.00731)	(0.00572)	(0.0194)
NUTS_5	3.578***	3.065***	0.972***	1.338***	1.803***	0.351***	1.062***	3.423***
	(0.0888)	(0.0489)	(0.0568)	(0.0321)	(0.0381)	(0.0432)	(0.0377)	(0.0682)
NUTS_3	1.491***	1.191***	0.314***	0.654***	0.645***	0.298***	0.374***	1.295***
	(0.0736)	(0.0374)	(0.0268)	(0.0179)	(0.0219)	(0.0223)	(0.0185)	(0.0569)
NUTS_2	0.222***	0.244***	0.00611	0.154***	0.0714***	0.0774***	-0.0586***	0.00105
	(0.0683)	(0.0338)	(0.0259)	(0.0167)	(0.0198)	(0.0209)	(0.0163)	(0.0501)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Time-Varying F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination Time-Varying F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	195,026	195,026	195,026	195,026	195,026	195,026	195,026	195,026
R-squared	0.889	0.873	0.119	0.502	0.416	0.105	0.197	0.829

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 2: *Fixed effects estimation with GTC.*

For the regressions of the extensive margin decomposition, the number of shipments per commodity (frequency) and the number of different commodities shipped explains approximately the same proportion of the aggregate extensive margin for all administrative boundaries; particularly for the municipal border (*NUTS_5*) which shows a greater agglomeration of trade flows. The *NUTS_3* and *NUTS_2* variables reflect the same pattern as for the same first level decomposition; i.e., provinces (*NUTS_3*) reduce its importance as trade border while regions (*NUTS_2*) have not an important impact on trade flows. For the intensive margin, the coefficient associated to average tons (physical quantities) is the most relevant explaining this margin. Average tons shipped within the municipality are higher than inter-municipalities ones, although showing a decreasing trend between borders, and even exhibiting a negative impact at the regional level (*NUTS_2*); i.e., the average tons of inter-regional flows are higher than intra-regional ones, and therefore two municipalities situated in two different provinces within the same region will trade less than two municipalities situated in different regions. Also, thanks to the inclusion of total trade in physical quantities, we confirm the robustness of the coefficients when considering monetary measures as in the intensive margin; especially for the regional dummy (*NUTS_2*), which shows a non-significant sign.

Looking at the coefficients corresponding to transport costs, they present the expected signs in all trade decomposition variables, indicating the existence of increasing returns in transport. All estimations are robust to the three measures of transport costs, being travel time (Table 10 in the A) the most penalizing cost as it is related to road's accessibility and efficiency. As for the *Contiguity* variable it is significant in all regressions; that is, contiguous municipalities trade more among themselves than with more distant municipalities, reinforcing the market area interpretation of the distribution of trade flows. Besides, the GDP by origin and by destination reflecting the market potential attracting forces are very significant as we are using a panel data including origin and destination time-varying fixed effects. Finally, note the good fit achieved by the model (R^2) in explaining total trade flows and the extensive margin, in contrast to the low values obtained by Hillberry and Hummels (2008), although smaller when explaining the intensive margin.

With this first set of estimations we conclude that the use of very detailed measures of trade flows and transport costs reduces the impact of higher regional boundaries indicating the existence of a weak internal border effect which is only relevant when we attend to the municipal level. This confirms the existence of the "illusory effect" at high regional boundaries once it is controlled for when low levels administrative limits are brought into the equation, indicating a non-disruption of the market for very large administrative levels, in contrast to other studies for the Spanish case (Requena and Llano (2010); Garmendia et al. (2012)) or the Chinese case (Poncet (2005)). Concretely, these findings shed light on the idea that trade flows tend to agglomerate in short distances, meanwhile their density reduce as we increase the distance, drawing a trade pattern which motivate the use of structural test as to explicitly determine the point in which trade agglomeration finished. Also this stresses the importance of disaggregating regionally the trade flows and the transport costs if we want to measure the "real" border effect between areas.

Finally, we are interested in studying the dynamics of the internal border effect. To achieve this goal, we estimate the same models resorting to a cross-section analysis instead of pooling the data in a whole panel database. Table 3 presents the results of regressing model (5) dividing the database by years 2003 (initial year) and 2007 (final year) but considering only the first level decomposition (1) of the total value of trade into the extensive and the intensive margins.[22] To summarize the output table we present only the results for the three administrative boundaries as we are interested on the dynamics of the border effect on the trade margins -nevertheless, although not included, the *contiguity* variable and the GDP by *origin* and by *destination* are highly significant.

All administrative boundaries and transport costs reflect the same pattern for the total value of trade, i.e., there exist a slowdown tendency between 2003 and 2007. This finding re-

[22]We perform cross-section regressions for the years 2003 and 2007, so we only include time-invariant fixed effects by origin and by destination.

GTC	Boundary	2003	2007
Total Value of Trade	NUTS 5	3.795***	2.835***
	NUTS 3	1.581***	1.079***
	NUTS 2	0.179*	−0.0313
Extensive Margin	NUTS 5	3.151***	2.555***
	NUTS 3	1.184***	0.906***
	NUTS 2	0.174*	−0.00365
Intensive Margin	NUTS 5	0.958***	0.841***
	NUTS 3	0.363***	0.291***
	NUTS 2	0.00161	0.0201
Distance	Boundary	2003	2007
Total Value of Trade	NUTS 5	3.851***	2.921***
	NUTS 3	1.560***	1.145***
	NUTS 2	0.146	−0.0215
Extensive Margin	NUTS 5	3.265***	2.748***
	NUTS 3	1.191***	1.055***
	NUTS 2	0.171***	−0.0997*
Intensive Margin	NUTS 5	0.878***	0.880***
	NUTS 3	0.269***	0.327***
	NUTS 2	−0.0937	0.0286
Time	Boundary	2003	2007
Total Value of Trade	NUTS 5	3.721***	2.931***
	NUTS 3	1.505***	1.166***
	NUTS 2	0.0834	−0.00232
Extensive Margin	NUTS 5	3.040***	2.645***
	NUTS 3	1.074***	0.986**
	NUTS 2	0.0645	0.0184
Intensive Margin	NUTS 5	0.936***	0.843***
	NUTS 3	0.341***	0.294***
	NUTS 2	-0.0277	0.0195

Table 3: *Cross-section regressions for 2003 and 2007 (First Level Decomposition)*

Transport Costs	Average 2003-2007	Mean by years 2003	2005	2007	Growth Rates
GTC	333.61	347.69	343.87	305.97	-12.00%
Distance	313.51	313.02	313.39	313.49	0.15%
Travel Time	287.19	290.21	286.58	284.16	-2.08%

Table 4: *Transport costs variation along the period.*

flects that the internal border effect is not constant along the years.[23] Indeed, administrative levels do not agglomerate the same amount of trade within themselves during all the years. In 2003, the *Nuts_5* is the one with the highest impact on trade flows, even larger than in the panel data regression. Meanwhile, it is confirmed that the *Nuts_2* level exhibits a not significant, or even negative, effect on trade flows, indicating again that regional borders are not as important as trade literature used to remark. The extensive margin shows, for the three measures of transport costs, the same pattern as the total value of trade, while the intensive margin does not change to a greater extent during the period. In this sense, we could conclude that the reduction in the effect of administrative boundaries on trade flows arises mainly from the extensive margin.

3.3. Analysis in growth rates

Once we have analyzed the response of trade flows to geographical frictions, we analyze the effect of transport costs on the trade flows and its margins along the years, in order to determine the most suitable measure of transport cost when panel-time series data is available. Table 4 shows the mean of each transport cost, on average for the period 2003-2007 and for the individual years 2003, 2005 and 2007, plus the growth rate experimented by each one of them. GTC is the cost measure with the highest variation, while distance and travel time have changed to a lesser extent.

By individual years, again the GTC has reduced more than the other transport costs measures. As a result, we should expect a higher impact of the GTC on trade flows along the period; i.e., distance and travel time have a lower variation during 2003-2007.[24]

To further stress this idea we perform a regression based on growth rates following equation (6) to study to what extent variability on trade flows (total value of trade) are explained by variability on transport costs. Indeed, and according to the previous Table 4, we should expect a more significant and negative impact of the GTC on the trade flows in

[23]We have performed a set of mean tests to analyze whether each administrative boundary are statistically different between 2003 and 2007. Thus, we reject the null hypotheses of equal coefficients.

[24]Note that, in contrast to other measures of the distance where it is a straight line between two points, our actual distance can vary from one year to the other as it is the distance travelled by trucks between two municipalities.

comparison to the rest of transport costs.

$$\Delta X_{ijt} = \beta_0 + \beta_1 \Delta cost_{ijt} + year + \eta_{it} + \eta_{jt} + \epsilon_{ijt} \quad (6)$$

Table 5 shows the results of regressing the growth rate of trade flows and its margins against the growth rate of each transport costs. We have performed separated regressions for each transport cost in which the endogenous and the exogenous variables are transformed in first differences. We also include year fixed effects and origin and destination invariant fixed effect to control for unobserved heterogeneity and the effect of business cycles.

		Variables in Growth Rates		
	Variables	Total Value	Extensive Margin	Intensive Margin
Costs in	*GTC*	-0.0296**	-0.0355**	-0.0275*
Growth	*Distance*	-0.00181	-0.00336	0.00433
Rates	*Travel Time*	0.00281	0.000761	0.00461

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 5: *Variation effect of transport costs on trade flows.*

As expected, distance and travel time have no significant effects either on trade flows variations (growth rates) or its margins, because their changes from one year to another are quite reduced. By contrast, the GTC has a negative and significant impact on the three endogenous variables, as contended. Indeed, a reduction in the GTC leads to an increase in trade flows; that is, only reductions in the shipment's aggregate economic cost, particularly those related to truck efficiency such as fuel as well as the reduction of salaries –both the main components of trucking operating costs– brings larger trade flows. In fact, GTCs capture both infrastructure and operating costs variability, which renders it as the as the best measure capturing the effect of transport cost changes in panel data studies. These results confirm the GTC as the only suitable measure of transport costs, while distance and time are only suitable proxies of trade costs for cross-section estimations of gravity equations, i.e., geographical *"stationary"* or constant measures of transport costs as distance and travel time –almost the only cost measures included in empirical trade studies– are not adequate.

4. Structural trade patterns: breakpoints analysis

The regressions estimated in the previous section are *naïve* in the sense that they simply provide an average estimate of the effect of transport costs on trade flows, qualified by the customary second order effect. In this section we systematically undertake a series of structural tests to determine the transportation cost thresholds or breakpoints that define relevant discontinuities in their effect on trade flows. With them, we split the database to regress the same gravity specification as in the first part but accounting for the different distance thresholds. Additionally, the empirical evidence on the border effect phenomenon usually gives a "general" value for the border effect, which is corrected by different misspecifications such as Anderson and van Wincoop (2003) or Santos Silva and Tenreyro (2006, 2011), among others. Nevertheless, it does not consider the possibility of having different border effects once we control for multiple distance thresholds. In this sense, Eaton and Kortum (2002) propose different intervals distances which act as trade barriers. These intervals are arbitrary as there is not an objective criterion to divide the transport cost data. The existence of at least one relevant threshold in the transport cost-trade flows relationship for the whole sample can be easily established at an approximate value of 150 euros by simple visual inspection of the kernel regressions: Over very short distances trade flows are radically affected by relative low transport costs (extensive margin), while for larger values are less dependent and becoming relatively "flat" (intensive margin). But further thresholds both against GTCs or its proxies can exist at different values. Indeed, the previous regressions confirm the idea according to which trade flows are mainly predetermined by the extensive margin which follows the same pattern on transport costs. As we understand it, when studying the magnitude and significance of the border effect, one should control for this non-linear relationship. Thus, we argue that one should test whether trade flows change radically once they have reached a given transport cost threshold, which we call breakpoint. That is, we argue that trade flows present different structural relationships (structural stability) with transport costs, resulting in relevant changes in the negative effect of administrative borders on trade. To determine these structural breakpoints we rely on Berthelemy and Varoudakis (1996), who present a Chow structural test for cross-section studies to check the structural stability of the parameters.[25] The test is carried out by establishing the model in which we want to determine the existence of a structural change and the threshold variable which creates these breakpoints. In our case, equation (5) represents the *baseline* model to develop the analysis fixing each transport cost as the threshold variable. However, we have not included the GDP by origin and by destination as trade flows present an increasing behaviour with these variables which would distort the regression and cause the failure of test. Additionally, we have performed the Chow test several times to check the existence of successive distance thresholds. To summarize the information, instead of performing the test for each year in the sample, we have chosen the mean of each variable in (5) to get single points for the whole period (2003-2007). Table 6 reports the breakpoints for the total value of trade and the extensive margin. Unsurprisingly, the test fails in detecting breakpoints for the intensive margin as its components do not drive reduction in trade as shown in the kernel regressions (Fig. 2c), and in sharp contrast with the total value of trade and its determinant, the extensive margin. Finally, all *breakpoints* obtained are significant at the

[25]This test divides the sample in two sub-samples. Afterwards, it begins to perform several iterations with the number of observations in each sub-sample changing continuously until it finds a breakpoint, in terms of the threshold GTC variable, in which the Chow test rejects the null hypotheses (at the 5% significance level) of the non-existence of a structural change in our model. This test presents the advantage that it is not necessary to provide an exogenous point in which the researcher suspects the possibility of a structural change, as in time-series models. By contrast, it endogenously determines this exact breakpoint in which the endogenous variable pattern changes significantly. This test has been recently implemented in Diallo (2012).

Total Value of Trade	Break Points obtained for the period 2003-2007				
	Break Point 1	Break Point 2	Break Point 3	Break Point 4	Break Point5
GTC (Euros)	189***	233***	285***	513*	706***
Distance (Km)	151***	202***	248***	351***	620***
Time (Min)	103*	139***	178***	240***	509***
Extensive Margin	Break Point 1	Break Point 2	Break Point 3	Break Point 4	Break Point5
GTC (Euros)	185***	246***	321***	582***	655***
Distance (Km)	151***	202***	318***	439***	761***
Time (Min)	103***	138***	180***	423***	528***

Robust standard errors. Standard errors in parentheses. Significance level: *** p<0.05, *p<0.3

Table 6: *Break Points for each Transport Cost.*

5%, except otherwise indicated.

Table 6 reinforces Hillberry and Hummels (2008) findings and qualifies them as we consider the full spectrum of administrative borders, whose presence may not exist at each range of trade values (e.g., intra-municipal trade at a Nuts-5 level cannot be normally observed for shipments with transport costs over the first breakpoint corresponding to 189 euros, since no municipality is so wide geographically). These results provide strong evidence that the trade flows are caused by the extensive margin as the first and the second breakpoints are mostly the same as those obtained for the total trade flows, and considering either the GTC or its proxies. Additionally, trade flows are highly concentrated at low transport cost values (around 189 and 233 euros). But after these two thresholds the difference between breakpoints becomes larger, indicating a declining tendency of trade flows on transport costs. The three measures of transport costs present the same pattern as those of the GTCs as they are highly correlated in cross-section estimations (Table 1). Moreover, from the transport cost database it is possible to relate travel times with distance by considering the legal speed limit: 90 km/hour, while the relation between GTCs and distance is about 1.1€/km. With these equivalences we could approximate the relation between transport costs, obtaining distances values similar to those shown by the breakpoints. However, this general relationship between the three cost magnitudes is about true for the breakpoints *1* through *4*, but not for the final breakpoint as the mandatory stop regulation (every 4h 30' the driver must rest 45') results in discontinuities in time and GTC, explaining why breakpoint *5* for time and GTCs is larger than their associated distance.[26]

To show further evidence in favor of the breakpoints obtained by way of the Chow test and stressing their consistency with our trade database, Figure 5 reflects the first breakpoint

[26]In this sense, we would subtract 45 minutes to the travel time for every four hours to get the "real" equivalency with the distance.

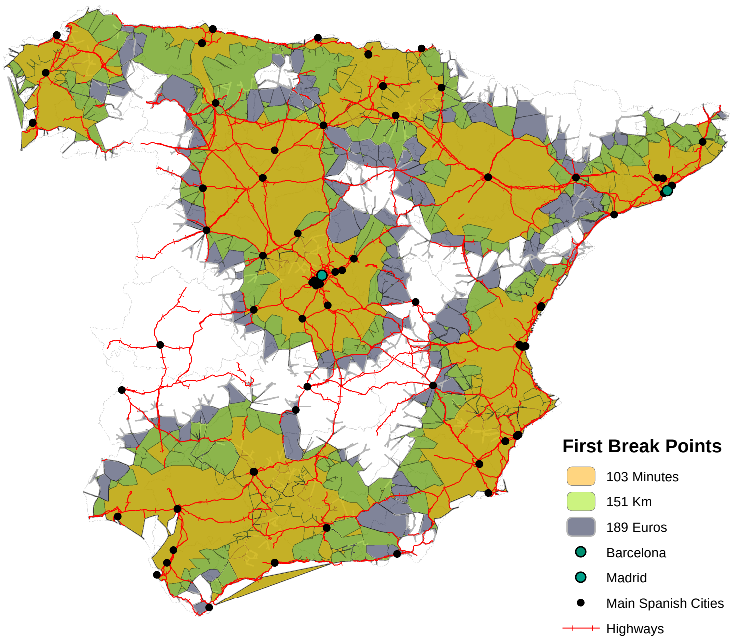


Figure 5: *First breakpoint obtained by the Chow test: Distance, Travel Time and GTC. Mean for the whole period data 2003-2007.*

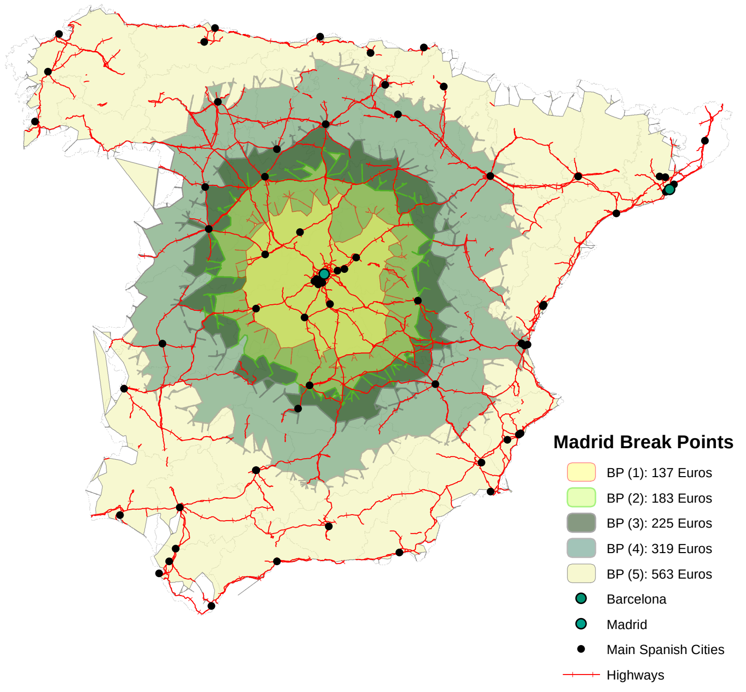
for the three measures of transport costs taking into account the 25 most populated cities. This map presents a first indication of *natural trade areas* in terms of the transport reach along the existing road network. The Arc/GIS Network Toolbox allows us to calculate the exact coordinates corresponding to the maximum distance given the type of road (arc) and its specific attributes as capacity, gradient, congestion level, etc. It can be observed that cities with a high volume of surrounding highways exhibit a larger trading area; i.e., each breakpoint represents an isocost line in terms of GTCs, distance and time. Indeed, those municipalities whose high capacity road infrastructure endowment are relatively scarce, or are surrounded by natural barriers (as cities such as Bilbao, Santander or Oviedo in northern Spain) are penalized in their accessibility, and their trade areas radii of influence are shorter. On the contrary, cities located in plateau areas (the Iberian Peninsula *mesetas*) such as Madrid, Zaragoza or Valladolid show longer accessibility. Depending on the transport cost, the influence area is larger or shorter. For the GTC and distance, being the less penalizing transport costs according to the first set of regressions reported in Table 2 and Table 9 in Appendix A on page 37, their area of influence is larger, while for travel time it is more limited. Additionally, as presented in Figure 5, the most populated cities with higher GDPs have market areas that encompass other main Spanish cities. It would be indicating that the most populated cities act as supply centers for the rest of the smaller cities. These areas overlap only if the cities are close enough; i.e., they can serve cities situated under the shadow of two large cities competing for the firms and consumers within it. However, for most of the cities it is observed that they tend to define a natural trade area which approximately ends where the other starts, and therefore they do not overlap. We term these market areas as natural trade areas, because they have been obtained by way of the objective procedure represented by Chows structural test, and taking into account only to trade flows and geographical frictions (trade costs).

In a further step Figures 6a and 6b shows the breakpoints for the two largest Spanish cities: Madrid and Barcelona, respectively.[27] Observing these two maps multiple empirical aspects can be remarked: i) the first breakpoint refers to the supply center between the city and its metropolitan area and other small cities not far from it; ii) the second breakpoint reaches some important cities (provincial capitals), especially from Madrid; iii) the third point presents the same pattern as the previous one, although it links to higher and richer cities such as Valladolid, Burgos and Salamanca for the case of Madrid; iv) the fourth point appears as very relevant as it joints Madrid and Barcelona with the richest cities in Spain in terms of trade, these are mainly Valencia, Zaragoza; v) the last breakpoint joins Madrid and Barcelona, indicating that trade flows within Spain overlap by these two cities; vi) the road network centrality is highly important as Madrid's trade area is always larger than Barcelona's for all the breakpoints; Madrid almost reaches all the Spanish regions within the last breakpoint, while Barcelona leaves half of the Iberian Peninsula out of its direct reach.

With these findings we argue that the methodology based on the Chow structural test represents a promising procedure to determine trade areas because, as far as we know from the literature on market areas, there have not been explicit empirical measurements of cities market area—Löffler (1998). Specifically, these empirical regularities are in line with the intuition proposed by the Lösch and the Christaller's model within the theoretical framework called the Urban Hierarchical System and the Central Place Theory (CPT)—McCann (2001); Parr (2002); Tabuchi and Thisse (2011); Mulligan et al. (2012). According to this framework, cities have market areas that are decreasing on transport costs, and where the largest cities producing diversified goods under increasing returns to scale can reach the furthest locations, meanwhile smaller cities have a reduced influence because they provide more standardized goods normally characterized by constant returns to scale. Although these ideas will be further explained in the next section, we now anticipate that, as predicted by these two models, these market areas represent geographical locations where cities compete with each other, with the competition reaching longer radii the higher the city-ranks.

Once we have determined the exact breakpoints where the effect of each transport cost on trade flows change, we split our database according to these thresholds. We perform a set of regressions using equation (5) but, as anticipated, we have eliminated the administrative boundary for the interval in which, on average, we should not expect any observation. As a clear example, for a transport cost interval of 285-513 euros (breakpoints *3* and *4*) there

[27]The GTC presented are in terms of distance equivalency as to correctly measure the market area on the Arc/GIS software.



(b) *Barcelona*

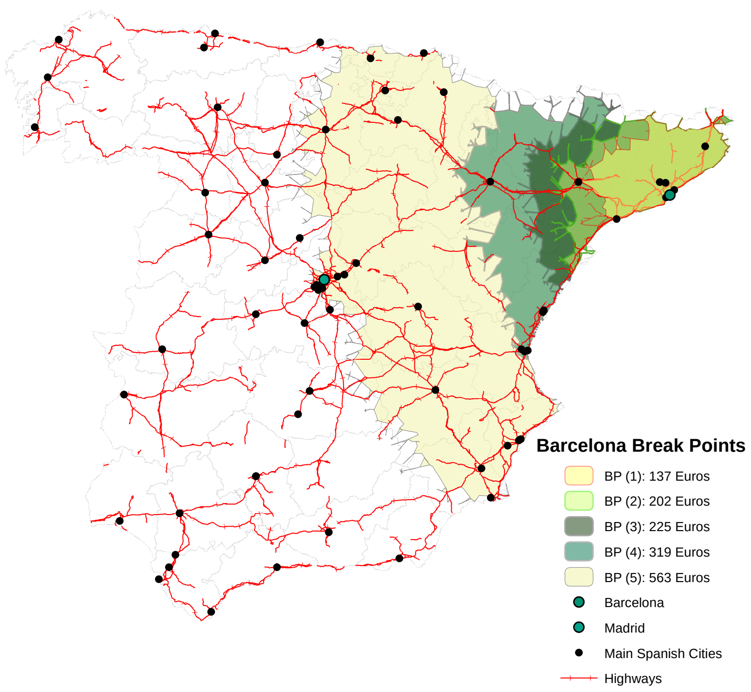


Figure 6: *Natural trade areas using distance.Breakpoints using the whole period data 2003-2007.*

are no shipments either at the intramunicipal level or between contiguous municipalities; and therefore we drop these variables in the regressions for this interval. We follow the same rationale for the rest of transport costs intervals. In fact, we contend that this is the correct methodology for measuring the internal border effect when accounting for the precise transport cost-trade value (non-linear) relationship; otherwise we would obtain an overall border effect which is not *real* in the sense that it would not attend to the specific characteristic of each shipment. To demonstrate this idea, we include in the same regressions a "general specification" (naïve gravity) using again equation (5) but without controlling for distance thresholds, and resorting to the quadratic distance as a measure of non-linear effect of distance on trade flows. As we show in what follows, administrative boundaries overestimate the effect of the border effect on short distances, meanwhile it has different effect when we split the distance by thresholds. Table 7 on the next page shows the results for the fixed effects estimation by GTCs thresholds for the total value of trade (upper panel), the extensive margin (middle panel) and the intensive margin (lower panel)–A reports equivalent regressions for distance and travel time. The GTC is highly penalizing on short distances, while it reduces its negative impact on trade flows, being even not significant for intermediate distances. GDPs by origin and by destination are introduced in levels to not eliminate non-linear relations with the trade flows specially on short distances. GDPs are highly significant for all thresholds. This indicates that transport costs are not as detrimental to trade as normally thought, as long as GDPs are big enough to attract trade flows, i.e., the classic trade-off of the gravity equation. What is interesting is that the relative pull of the GDPs is generally larger for destination GDPs, particularly for very short transport cost values. Administrative boundaries are very penalizing on the first threshold, especially the municipal border as in the nave regression for the whole sample in Table 2. *NUTS_2* is only significant for longer distances although it disappears for very long distances. It is non-significant for intermediate distances as the *NUTS_3* level captures the border effect. We have performed several analyses to confirm this idea. Finally, all coefficients are distorted or even overestimated if we do not control for the distance thresholds.

The extensive margin presents the same decreasing pattern on distance, being even non-significant for the interval 582-655 euros. Here, GDPs are always positive and significant and the administrative boundaries present a positive effect on trade. As in the regression for the total value, *Nuts_3* has an increasing impact on trade flows, while *Nuts_2* is only significant for the fourth threshold. In the three regressions, the model presents a good fit (high R^2) on shorter transport costs although it is decreasing in their value, indicating that there exist other factors affecting trade flows to control for in the larger GTCs values. Indeed, we find clear evidence that the general regression is driven by the specification on short GTC values, distances and travel times; that is, the bulge of trade flows take place within short transport costs values. That is why we conclude that the border effect calculated on the trade literature is biased in the sense that it does not control for very detailed trade flows and transport costs. Also, this high density of trade areas within short transport cost values corroborate the existence of trade areas defined by these value thresholds, whose existence is explained in the literature by agglomeration economies, either external or internal to the firm.

(a) *Total Value of Trade*

	GTC (0 Euros - 189 Euros)	*GTC (189 Euros - 233 Euros)*	*GTC (233 Euros - 285 Euros)*	*GTC (285 Euros - 513 Euros)*	*GTC (513 Euros - 706 Euros)*	*GTC (More than 706 Euros)*	*General Specification*
Variables	**Total Value**	**Total Value**	**Total Value**	**Total Value**	**Total Value**	**Total Value**	**Total Value**
GTC	-0.0145*** (0.000759)	0.00265 (0.00294)	-0.00420 (0.00353)	-0.00244*** (0.000413)	-0.00324*** (0.000689)	-0.00141*** (0.000211)	-0.00529*** (0.000297)
GTC Square	---	---	---	---	---	---	2.86e-06*** (1.87e-07)
Contiguity	0.911*** (0.0748)	---	---	---	---	---	1.311*** (0.0699)
GDP$_i$	1.88e-08*** (6.86e-09)	2.28e-08*** (4.40e-09)	3.96e-08*** (6.45e-09)	2.34e-08*** (4.06e-09)	3.12e-08*** (5.37e-09)	1.77e-08*** (3.64e-09)	2.04e-08*** (2.67e-09)
GDP$_j$	3.26e-08*** (6.72e-09)	3.82e-08*** (3.71e-09)	3.46e-08*** (3.66e-09)	3.19e-08*** (7.27e-09)	2.92e-08*** (3.48e-09)	2.54e-08*** (5.69e-09)	2.51e-08*** (2.34e-09)
NUTS_5	3.022*** (0.0942)	---	---	---	---	---	3.296*** (0.0848)
NUTS_3	1.059*** (0.0777)	1.859*** (0.358)	2.678*** (0.429)	---	---	---	1.297*** (0.0648)
NUTS_2	0.427*** (0.0836)	0.0797 (0.127)	0.0168 (0.195)	0.189 (0.158)	1.358*** (0.229)	---	0.106 (0.0692)
Observations	17,532	2,421	2,072	7,914	4,720	4,386	39,045
R-squared	0.957	0.935	0.772	0.675	0.825	0.639	0.941
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(b) *Extensive Margin*

	GTC (0 Euros - 185 Euros)	*GTC (185 Euros - 246 Euros)*	*GTC (246 Euros - 321 Euros)*	*GTC (321 Euros - 582 Euros)*	*GTC (582 Euros - 655 Euros)*	*GTC (More than 655 Euros)*	*General Specification*
Variables	**Shipments**	**Shipments**	**Shipments**	**Shipments**	**Shipments**	**Shipments**	**Shipments**
GTC	-0.0140*** (0.000446)	-0.00357*** (0.00128)	-0.00528*** (0.00125)	-0.00175*** (0.000319)	-0.00231 (0.00192)	-0.00113*** (0.000148)	-0.00665*** (0.000212)
GTC Square	---	---	---	---	---	---	3.62e-06*** (1.32e-07)
Contiguity	0.767*** (0.0383)	---	---	---	---	---	1.126*** (0.0353)
GDP$_i$	3.06e-08*** (7.25e-09)	2.47e-08*** (4.01e-09)	1.86e-08*** (4.03e-09)	2.01e-08*** (4.03e-09)	3.41e-08*** (3.89e-09)	1.52e-08*** (3.67e-09)	2.35e-08*** (4.15e-09)
GDP$_j$	2.60e-08*** (7.42e-09)	3.65e-08*** (2.49e-09)	3.84e-08*** (3.04e-09)	2.49e-08*** (5.02e-09)	1.74e-08*** (3.15e-09)	2.04e-08*** (3.99e-09)	2.60e-08*** (5.40e-09)
NUTS_5	2.556*** (0.0519)	---	---	---	---	---	2.867*** (0.0533)
NUTS_3	0.808*** (0.0443)	1.276*** (0.176)	1.553*** (0.226)	---	---	---	1.083*** (0.0437)
NUTS_2	0.293*** (0.0516)	-0.0164 (0.0811)	0.161 (0.109)	0.553*** (0.129)	---	---	0.112*** (0.0423)
Observations	17,316	3,172	2,860	8,607	1,834	5,256	39,045
R-squared	0.959	0.880	0.809	0.631	0.896	0.718	0.950
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(c) *Intensive Margin*

	GTC (0 Euros - 189 Euros)	*GTC (189 Euros - 285 Euros)*	*GTC (285 Euros - 513Euros)*	*GTC (513 Euros - 706 Euros)*	*GTC (More than 706 Euros)*	*General Specification*
Variables	**Average Value**	**Average Value**	**Average Value**	**Average Value**	**Average Value**	**Average Value**
GTC	-0.00422*** (0.000334)	8.78e-05 (0.000831)	-0.000510** (0.000223)	-0.000855** (0.000434)	-0.000962*** (0.000113)	-0.000413*** (8.82e-05)
GTC Square	---	---	---	---	---	1.62e-07*** (5.57e-08)
Contiguity	0.286*** (0.0343)	---	---	---	---	0.463*** (0.0312)
GDP$_i$	5.21e-09 (3.24e-09)	1.21e-08*** (2.32e-09)	1.05e-08*** (2.50e-09)	1.28e-08*** (3.11e-09)	4.09e-09 (2.75e-09)	9.03e-09*** (1.36e-09)
GDP$_j$	9.34e-09*** (3.25e-09)	0.608*** (0.170)	1.42e-08*** (4.61e-09)	1.53e-08*** (1.54e-09)	1.12e-08** (4.71e-09)	1.05e-08*** (1.49e-09)
NUTS_5	0.847*** (0.0568)	---	---	---	---	0.998*** (0.0609)
NUTS_3	0.257*** (0.0359)	0.150* (0.0781)	---	---	---	0.336*** (0.0286)
NUTS_2	0.138*** (0.0369)	1.235*** (0.250)	0.0460 (0.104)	0.114 (0.150)	---	0.0555** (0.0270)
Observations	17,532	4,493	7,914	4,720	4,386	39,045
R-squared	0.406	0.575	0.387	0.445	0.451	0.290
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 7: *Fixed effects estimation by GTC breakpoints.*

Finally, the intensive margin presents the same pattern as for the total value of trade as well. In this case, we have aggregated two distance intervals (233–285 euros and 285–513 euros) in order to achieve results due to the few degrees of freedom obtained once we control for origin and destination fixed effects. *NUTS_3* and *NUTS_2* are only significant for low and intermediate transport cost values while *NUTS_5*, although positive, would result in an overestimated border effect if we do not control for distance thresholds. Again, GDPs are highly significant, although for lower and middle GTCs only the destination GDP is consistently significant. This indicating that destinations has a greater effect on trade; that is, even the GTCs is not as determinant on trade provided that the destination GDP is high enough to attract trade flows.

5. Trade areas and the hierarchy of Spanish cities

The breakpoint thresholds previously calculated convey relevant empirical findings in terms of the central place theory and its associated hierarchical urban system. Following this literature based on the Lösch and Christaller's model (McCann (2001); Parr (2002); Mulligan et al. (2012)), we would expect areas of influence whose geographical reach is driven by transport costs; that is, consumers and firms will locate in places where they can be supplied by different cities, and taking into account the transport costs in which they incur because of their consumption or production processes. In the model, cities cover all locations as long as consumers and firms are willing to cover the transports costs of having the goods shipped to their particular location. As a result of this demand schedule, cities market areas show a decreasing behavior in shipments as transport costs increase, eventually coming into spatial competition with other cities for the geographical space where their areas of influence overlap. To show this idea, Figure 16 in Appendix D shows the kernel distribution of the two largest Spanish cities: Madrid and Barcelona. It is clear that there exists spatial competition between both cities, although they present interesting differences in their trade patterns. While Barcelona spreads along the geography, i.e., it presents a larger trade area, Madrid shows a higher density of trade in shorts distances. That is because the metropolitan area of both cities leads them to supply the surrounding cities in different ways. To meet intermediate production processes and final demand, Madrid has to reach nearby large cities with a high volume of trade by road, while Barcelona uses its port as a supplying hub for its surrounding cities.

In a stylized and simple version of the model these market areas are spatially represented by nested hexagons within a geographical lattice, where the city is in the middle of each hexagon and its radii of influence are increasing in city's size. It means that when the radii reach a specific point (breakpoint in our previous analyses), the market area changes pattern radically until it eventually disappears for the case of smaller cities. Specifically, an urban hierarchy à la Christaller emerges because considering a range of different goods, few cities can serve high-order items (specialized goods) which results in larger GTCs thresholds. Meanwhile, as the order-scale of items reduce (low-order items), smaller cities provide more standardized products. In this sense, *"the most central location in the entire system provides all of its goods and services thereby satisfying the so called exhaustive principle. But, moving* *down through this functional continuum (of goods), other locations on the landscape are sufficiently well located to provide some, but not all, goods that are provided at the most central location"* (Mulligan et al. (2012), pp. 404). So, according to Christaller's model we should expect that, within a country, a hierarchical system of cities emerges where few cities (*Rank 1 cities* or *Dominant cities*) present the largest market areas supplying the full range of products; a second group of big cities (*Rank 2 cities*) serves a huge variety of commodities with a large market area too; while in the next levels other medium size and small cities (*Rank 3 cities* and *Rank 4 cities* respectively) are scattered geographically between these two previous city-ranks, presenting more standardized products with a lower or even insignificant market areas, and supplying the most homogenous goods needed for consumption.

Although the evidence confirming the existence of the predicted urban hierarchy system is related to the magnitude of trade flows and its associated product mix, empirical studies have focused on a proxy corresponding to cities populations, showing that the larger the population, the higher will be its rank within the country. The underlying assumption is that population size is not only a good proxy of city's size, but also implies a more diversified demand that supports the production of a wider range of products. Despite this attempt to relate the city's size with its market area, we consider that there is lack of empirical evidence based on relevant trade flows. That is why, so as to complement and reinforce our breakpoint thresholds and border effect analyses, we rely on our empirical data to provide evidence that supports the existence of an urban system based on the volume and sectoral characteristics of the distribution of trade flows, i.e., the number of commodities and trading partners that each city shows, and its relation with the city's market area.

Table 8 shows these distributions in 2003 and 2007, differentiating by intervals the number of commodities shipped and the numbers of regions (municipalities) with which shipments take place. Data are given as percentage over the total number of municipalities, indicating the amount of municipalities which trade a determined number of commodities and the number of different municipalities with which they trade.[28] It can be observed that the largest number of municipalities trade between 10 to 50 commodities with a set of 10 to 50 municipalities interval, although this percentage has shifted in 2007 as the number of commodities shipped to the same number of municipalities increases, showing a diversification in the shipments' product mix–the shipments' product composition, as well as a wider array of destinations. In fact the percentage of shipments in the upper intervals, greater than 10 commodities and 10 municipalities, increase from 87.86% in 2003 to 91.55% in 2007. Finally, there has been an increase in the number of municipalities that trade more than 100 commodities to more than 100 regions: from 14.19% to 15.50%, respectively.

Table 8 sheds light on the urban hierarchy system. As it can be observed, the main diagonal characterizes different types of cities in the lines expressed above. Specifically, we have clustered the municipalities according to the two variables, the number of different commodities that they ship and the number of trading partners. We have used cluster tech-

[28]The intervals considered are in line with Mayer and Ottaviano (2008) for the case of exporting firms, but adding some intervals to better disentangle the regional and sectorial shipments distributions.

2003	*Number of Municipalities*					
Number of commodities	1	(1-5]	(5-10]	(10-50]	(50-100]	More than 100
1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
(1-5]	0.00%	0.16%	1.12%	3.19%	0.00%	0.00%
(5-10]	0.00%	0.32%	1.44%	5.42%	0.16%	0.00%
(10-50]	0.00%	0.00%	0.32%	38.6%	11.0%	0.16%
(50-100]	0.00%	0.00%	0.00%	1.91%	17.3%	2.71%
More than 100	0.00%	0.00%	0.00%	0.00%	1.91%	14.19%
2007	*Number of Municipalities*					
Number of commodities	1	(1-5]	(5-10]	(10-50]	(50-100]	More than 100
1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
(1-5]	0.00%	0.16%	1.12%	1.60%	0.00%	0.00%
(5-10]	0.00%	0.00%	0.64%	3.83%	0.00%	0.00%
(10-50]	0.00%	0.00%	1.12%	42.1%	8.31%	0.00%
(50-100]	0.00%	0.00%	0.00%	1.76%	19.1%	1.60%
More than 100	0.00%	0.00%	0.00%	0.00%	3.04%	15.5%

Source: Own elaboration from the Road Freight Transportation Survey data.

Table 8: *Shipments distribution by municipalities and products.*

niques (K-means based on centroid distances) obtaining four different groups corresponding to Christaller's idea of product diversity and geographical reach/rank.[29]

According to the theory, each of these city-groups should present a different market area that reduces with the rank of the city. Fig.7 depicts the specific relationship by showing different kernel regressions for each cluster of cities. As expected, cities of a higher rank present higher densities for all distances, thereby enveloping the distributions of lower rank cities-both at any trade cost value and threshold. Moreover, the elasticity of shipments to GTCs is lower for all transport costs; i.e., in the short distances *Rank 3* and *Rank 4* cities are not as sensitive to GTCs as higher rank cites. Indeed, the agglomeration of trade flows on short distance observed in Figs.2a-c is clearly driven by these two city-clusters.

We map this information to show a detail representation of the Spanish urban hierarchical system. As far as we know Figure 8 on page 33 represents the first illustration of an actual hierarchical system based on trade flows. It illustrates cities in relation to the four city-type clusters already calculated. Additionally, we map the breakpoint thresholds for the dominant (*Rank 1*) cities after applying again the Chow test only for this group of four cities (Madrid, Barcelona, Valencia and Seville). It can be observed that the first threshold covers the

[29]It is remarkable that this hierarchy emerges from the raw data; i.e., based on the dendogram we decided to form four groups of cities, and using the K-means analysis in a subsequent step, obtained city-type clusters that exceptionally match the real economic distribution of Spanish cities.

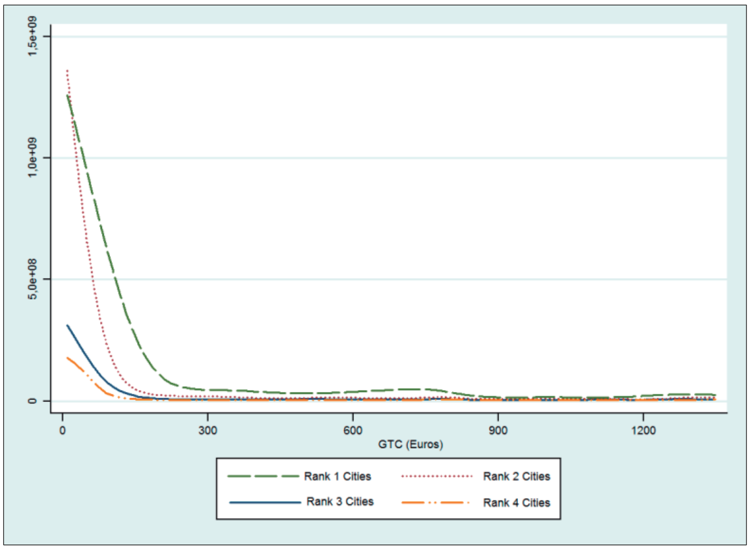


Figure 7: *Kernel regressions for different city-cluster.*

metropolitan area of these big cities, while the second and the third breakpoints reach *Rank 2* cities or even intermediate (*Rank 3*) cities as predicted by the central place theory. Indeed, it seems that the map is a remarkable representation of the geometry proposed by Christaller (Parr (2002)); especially, if we consider that his simplest theoretical representation was based on a homogeneous space with no transport costs. Moreover, it is clear that the geographical reach of these breakpoints exceeds all administrative levels of the spatial units; particularly, Nuts-5 (municipal) and Nuts-3 (provincial) territorial units. Also, there exists an additional breakpoint at the value of 665€, although we have not shown it to summarize the map information. This last breakpoint represents again the distance between the main *Rank 1* cities, particularly Madrid and Barcelona, which emphasizes the results obtained in the previous section.[30] This pattern explains the high border effect found at the municipal level in Table 2, indicating that it is the trade flows between first rank cities what shapes the structural trade distribution of the data, thereby resulting in high market areas, i.e., large breakpoints. In fact, and taking into account that each successive breakpoint entails lower levels of trade flows, this result indicates that the border effect within a country is due to agglomeration economies, because the farther away we move from larger cities, the lower the negative effect on trade flows, and the lower the border effect is.

6. Conclusions

In this study we have analyzed the structure of the internal border effect within Spain and the role played by transport costs in hampering trade flows, using two novel micro databases

[30] Additionally, calculating the breakpoints for Rank 2 cities we obtain a third point that exactly matches the same GTC value (261€), of that obtained for *Rank 1* or *Dominant cities*. This finding emphasizes the idea of competing market areas McCann (2005).

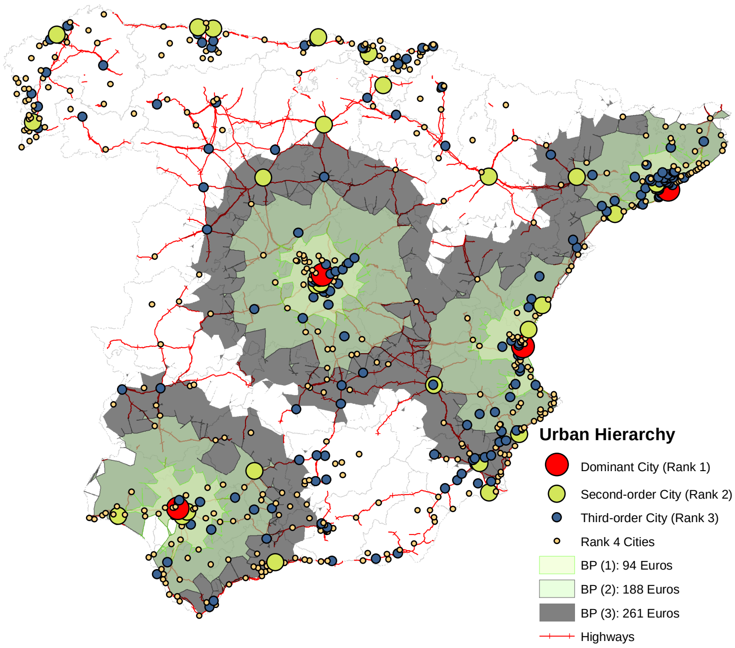


Figure 8: *Urban Hierarchy and Natural Trade Areas for Dominant Cities. First, Second and Third GTCs Breakpoints.*[a]

[a]The fourth breakpoint is equal to 665€.

in the literature on interregional commerce. The first one compiles information about freight shipments transported by trucks between Spanish municipalities for the period 2003-2007. The second one relies on a very precise measure of transport cost (the generalized transport cost, GTC), and calculates actual trade cost—monetary—values that comprise both distance and time economic costs associated to each optimal route between cities. Indeed, we show that this GTC measure is the most suitable variable when explaining trade flows within a panel data structure–in contrast to distance and travel time normally used in the literature. This is because GTC is the only measure capturing the real dynamics and effects on trade flows of changes in operating costs over the years. Thanks to this detailed information on shipments we are able to decompose aggregate trade flows into their extensive and intensive margin, so as to determine the effects of the geographical frictions corresponding to the three territorial boundaries existing in Spain (home bias—or border—effects) on each one of them, while controlling for the GTC measure. In a first *naïve* analysis with all trade data that does not allow for a precise identification of the trade flow-transport cost thresholds, we conclude that the internal border effect varies in sign and magnitude depending on the administrative boundary and by each margin. Specifically, the results on the effects of the internal administrative levels on trade flows are higher than the ones reported by

Hillberry and Hummels (2008) for low transport cost values. Moreover, for all cost measures studied, a decreasing impact of the border effect on trade is observed as goods are shipped to destinations located in regions different from their own region of origin. In this sense, regional borders have a much reduced, or even negative, impact on the trade flows taking place within them; while intra-municipal trade (or surrounding areas) tends to agglomerate the largest share of trade flows.

In contrast to Hillberry and Hummels (2008), which suggest that this city's border effect is a *reductio ad absurdum* in the "home bias" literature, we argue that this higher density has nothing to do with border impediments, but it arises as a result of transport cost thresholds that define the geographical reach of agglomeration economies, which mass shipments around *rank 1* or *dominant cities* resulting in large market areas. To show this idea we introduce the endogenous Chow test into the trade literature, allowing us to determine the transport costs thresholds at which the trade flows–transport cost relationship changes structurally. Thanks to this methodology, we run equivalent regressions for each range of shipments within subsequent breakpoints, so as to correctly measure the border effect within a country and define cities' market areas. In this respect we confirm that the internal border effect is not present for shipments between the usual range corresponding to Nuts-2 regions, as they take at shorter distance. The high density of shipments at GTC values below 300 euros presents strong evidence on how agglomeration economies shape market areas and an urban hierarchy based on trade flows, where few large cities dominate geographically and account for the largest trade share, while small cities spread along the geography without significant influential areas in terms of market size.

We argue that these results are in line with the empirical predictions emanating from Central Place Theory and produce the first map of an urban hierarchy based on actual trade flows, including the corresponding regional boundaries between cities. Moreover, we provide evidence corroborating the hypothesis that associates the ranking of the cities not only to trade volumes, but most importantly to the diversity of the goods being shipped (production mix which passes on to exports) and their multiple destination geographical reach. All these results call for future studies based on how this very sharp picture of trade flows and market areas responds to the existence of an urban hierarchy. In this sense, it is necessary to expand the analysis focusing on the intensive margin (average value of each shipment) at the sectorial level, so as to understand in what sense it leads to the specialization of the economic structure of Spanish cities. Finally, focusing on sectorial analyses, it is worth explaining the existence of large trade flows between far away (*rank 1* or *dominant*) cities, which in turn implies the study of the value transport cost relative to the origin (mill) and destination process subject to high transport costs; i.e., a challenging analysis in terms of relative prices and demand behavior. Particularly, establishing whether goods that are shipped to farther locations are highly differentiated (heterogeneous) as opposed to goods that are traded over very short distances as they have close substitutes in other locations (homogenous goods).

Acknowledgments

We thank Thierry Mayer, Johannes Bröcker and, particularly, Horst Raff at the Kiel Institute for World Economy and Kiel University, and Geoffrey Hewings at R.E.A.L. (University of Illinois at Urbana-Champaign), for their useful comments. We also express our gratitude to Francisco Requena-Silvente (Universitàt de Valencia and Sheffield University), Jesús Fernández-Huertas Moraga (FEDEA) and Tamara de la Mata, IESE Business School.

We acknowledge support from research projects ECO2010-21643 financed by the Ministerio de Economía and the TRANSPORTRADE (S2007-HUM-0467) project financed by the Comunidad Autónoma de Madrid. Jorge Díaz-Lanchas acknowledges funds received from the FPU-2010 program and the National Prize IX Certámen Arquímedes for Young Researchers, both financed by the Ministerio de Educación.

7. References

Anas, A., Xiong, K., 2003. Intercity trade and the industrial diversification of cities. Journal of Urban Economics 54 (2), 258–276.

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

Baldwin, R., Taglioni, D., September 2006. Gravity for dummies and dummies for gravity equations. Working Paper 12516, National Bureau of Economic Research.

Behrens, K., 2005. Market size and industry location: traded vs non-traded goods. Journal of Urban Economics 58 (1), 24–44.

Behrens, K., Mion, G., Murata, Y., Südekum, J., 2013. Spatial frictions. Tech. rep., Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit.

Benedictis, L., Taglioni, D., 2011. The gravity model in international trade. In: De Benedictis, L., Salvatici, L. (Eds.), The Trade Impact of European Union Preferential Policies. Springer Berlin Heidelberg, pp. 55–89.

Berthelemy, J., Varoudakis, A., April 1996. Economic growth, convergence clubs, and the role of financial development. Oxford Economic Papers-New Series 48 (2), 300–328.

Borraz, F., Cavallo, A., Rigobon, R., Zipitría, L., 2012. Distance and political boundaries: Estimating border effects under inequality constraints. Tech. rep., National Bureau of Economic Research.

Brülhart, M., Trionfetti, F., 2009. A test of trade theories when expenditure is home biased. European Economic Review 53 (7), elsevier.

Cavailhes, J., Gaigne, C., Tabuchi, T., Thisse, J.-F., 2007. Trade and the structure of cities. Journal of Urban Economics 62 (3), 383–404.

Chen, N., 2004. Intra-national versus international trade in the european union: why do national borders matter? Journal of International Economics 63 (1), 93 – 118.

Combes, P.-P., Lafourcade, M., Mayer, T., 2005. The trade-creating effects of business and social networks: evidence from france. Journal of International Economics 66 (1), 1 – 29.

Diallo, I. A., 2012. Suchowtest: Stata module to calculate successive chow tests on cross section data.

Duranton, G., Morrow, P., Turner, M. A., 2013. Roads and Trade: Evidence from the US. Centre for Economic Policy Research.

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

Evans, C. L., 2003. The economic significance of national border effects. The American Economic Review 93 (4), pp. 1291–1312.

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

Garmendia, A., Llano, C., Minondo, A., Requena, F., 2012. Networks and the disappearance of the intranational home bias. Economics Letters 116 (2), 178 – 182.

Head, K., Mayer, T., 2002. Illusory border effects: distance mismeasurement inflates estimates of home bias in trade. Vol. 1. CEPII Paris.

Hillberry, R., 2002b. Commodity flow survey data documentation. Tech. rep., U.S. International Trade Commission, Forum for Research in Empirical International Trade (F.R.E.I.T.).

Hillberry, R., Hummels, D., 2003. Intranational home bias: Some explanations. Review of Economics and Statistics 85 (4), 1089–1092.

Hillberry, R., Hummels, D., 2008. Trade responses to geographic frictions: A decomposition using micro-data. European Economic Review 52 (3), 527 – 550.

Hillberry, R. H., 2002a. Aggregation bias, compositional change, and the border effect. Canadian Journal of Economics/Revue canadienne d'conomique 35 (3), 517–530.

Llano, C., Esteban, A., Pérez, J., Pulido, A., 2010. Opening the interregional trade black box: The c-intereg database for the spanish economy (19952005). International Regional Science Review 33 (3), 302–337.

Llano-Verduras, C., Minondo, A., Requena-Silvente, F., 2011. Is the border effect an artefact of geographical aggregation? The World Economy 34 (10), 1771–1787.

Löffler, G., 1998. Market areas–a methodological reflection on their boundaries. GeoJournal 45 (4), 265–272.

Martin, W., Pham, C., 2008. Estimating the gravity model when zero trade flows are important.

Mayer, T., Ottaviano, G. I., 2008. The happy few: The internationalisation of european firms. Intereconomics 43 (3), 135–148.

McCann, P., 2001. Urban and regional economics. Vol. 15. Oxford University Press Oxford.

McCann, P., 2005. Transport costs and new economic geography. Journal of Economic Geography 5 (3), 305–318.

Melitz, M. J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.

Mulligan, G. F., Partridge, M. D., Carruthers, J. I., 2012. Central place theory and its reemergence in regional science. The Annals of Regional Science 48 (2), 405–431.

O'sullivan, A., 2003. Urban economics. McGraw-Hill/Irwin.

Parr, J. B., 2002. The location of economic activity: central place theory and the wider urban system. Edward Elgar Cheltenham.

Poncet, S., 2005. A fragmented china: measure and determinants of chinese domestic market disintegration. Review of International Economics 13 (3), 409–430.

Puga, D., 2010. The magnitude and causes of agglomeration economies. Journal of Regional Science 50 (1), 203–219.

Requena, F., Llano, C., 2010. The border effects in spain: an industry-level analysis. Empirica 37 (4), 455–476.

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

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

Santos Silva, J. M., Tenreyro, S., 2011. Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. Economics Letters 112 (2), 220–222.

Tabuchi, T., Thisse, J.-F., 2011. A new economic geography model of central places. Journal of Urban Economics 69 (2), 240–252.

Van der Marel, E., 2012. Determinants of comparative advantage in services. Tech. Rep. 87, FIW Working Paper.

Wei, S.-J., 1996. Intra-national versus international trade: how stubborn are nations in global integration? Tech. rep., National Bureau of Economic Research.

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

Yilmazkuday, H., 2012. Understanding interstate trade patte. Journal of International Economics 86, 158–166.

Zofío, J. L., Condeço-Melhorado, A. M., Maroto-Sánchez, A., Gutiérrez, J., Nov. 2011. Decomposing generalized transport costs using index numbers: A geographical analysis of economic and infrastructure fundamentals. Working Papers in Economic Theory 2011/06, Universidad Autnoma de Madrid (Spain), Department of Economic Analysis (Economic Theory and Economic History).

Appendices

A. Regressions considering Distance and Travel Time.

	First Trade Decomposition			Extensive Margin		Intensive Margin		
VARIABLES	Total Value T_{ij}	Number of Shipments N_{ij}	Average Value per Shipment $\overline{PQ}_{ij}$	Number of commodities N_{ij}^{K}	N. of shipments per commodity N_{ij}^{F}	Price $\bar{P}_{ij}$	Tons $\bar{Q}_{ij}$	Trade in Quantity Q_{ij}
Distance	-0.893***	-0.899***	-0.103***	-0.135***	-0.345***	-0.541***	-0.585***	-1.196***
	(0.052)	(0.0252)	(0.0152)	(0.0125)	(0.0111)	(0.0107)	(0.0139)	(0.0429)
Distance Sq.	0.206***	0.207***	0,000449	0.0411***	0.0748***	0.117***	0.145***	0.279***
	(0.0166)	(0.00991)	(0.00618)	(0.00448)	(0.005)	(0.00463)	(0.00725)	(0.0145)
Contiguity	1.332***	1.117***	0.461***	0.0933***	0.664***	0.648***	0.811***	1.240***
	(0.0636)	(0.0333)	(0.0305)	(0.0231)	(0.0229)	(0.0149)	(0.0263)	(0.0478)
log GDPi	0.362***	0.310***	0.215***	0.177***	0.133***	0.262***	0.111***	0.229***
	(0.0231)	(0.0144)	(0.0086)	(0.00776)	(0.00587)	(0.00523)	(0.00785)	(0.0197)
log GDPj	0.374***	0.348***	0.226***	0.203***	0.139***	0.264***	0.168***	0.262***
	(0.0242)	(0.0139	(0.00823)	(0.0073)	(0.00572)	(0.00518)	(0.00775)	(0.0191)
NUTS_5	3.292***	2.831***	0.924***	0.354***	0.939***	1.216***	1.655***	3.090***
	(0.0886)	(0.048)	(0.0576)	(0.044)	(0.0383)	(0.0316)	(0.0378)	(0.0668)
NUTS_3	1.196***	0.942***	0.265***	0.297***	0.242***	0.520***	0.483***	0.947***
	(0.0735)	(0.0361)	(0.0283)	(0.0238)	(0.0191)	(0.0172)	(0.0214)	(0.0543)
NUTS_2	0,0498	0.102***	-0,033	0.0804***	-0.146***	0.0692***	-0,029	-0.195***
	(0.0636)	(0.0321)	(0.0261)	(0.0214)	(0.0166)	(0.0159)	(0.0197)	(0.0498)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Time-Varying F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination Time-Varying F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	195.026	195.026	195.026	195.026	195.026	195.026	195.026	195.026
R-squared	0,89	0,874	0,119	0,104	0,198	0,507	0,418	0,83

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 9: *Fixed effects estimation with Distance.*

	First Trade Decomposition			Extensive Margin		Intensive Margin		
VARIABLES	Total Value T_{ij}	Number of Shipments N_{ij}	Average Value per Shipment $\overline{PQ}_{ij}$	Number of commodities N_{ij}^K	N. of shipments per commodity N_{ij}^F	Price $\bar{P}_{ij}$	Tons $\bar{Q}_{ij}$	Trade in Quantity Q_{ij}
Time	-0.935*** (0.0531)	-0.969*** (0.0296)	-0.0791*** (0.0183)	-0.187*** (0.0141)	-0.284*** (0.0132)	-0.584*** (0.0153)	-0.582*** (0.0186)	-1.195*** (0.0481)
Time Square	0.194*** (0.014)	0.202*** (0.00991)	0,000282 (0.00588)	0.0511*** (0.00413)	0.0536*** (0.00455)	0.124*** (0.00566)	0.130*** (0.00733)	0.243*** (0.0135)
Contiguity	1.378*** (0.063)	1.163*** (0.0332)	0.468*** (0.0303)	0.0975*** (0.0231)	0.695*** (0.0221)	0.677*** (0.015)	0.853*** (0.0264)	1.309*** (0.0479)
log GDPi	0.361*** (0.023)	0.309*** (0.0141)	0.214*** (0.0086)	0.178*** (0.00776)	0.130*** (0.00587)	0.261*** (0.00525)	0.109*** (0.00783)	0.226*** (0.0196
log GDPj	0.373*** (0.0241)	0.347*** (0.0136)	0.225*** (0.00824)	0.204*** (0.00731)	0.136*** (0.00572)	0.263*** (0.0052)	0.166*** (0.00775)	0.261*** (0.0191)
NUTS_5	3.562*** (0.0819)	3.070*** (0.0469)	0.982*** (0.057)	0.350*** (0.0428)	1.108*** (0.0378)	1.357*** (0.032)	1.849*** (0.038)	3.473*** (0.0646)
NUTS_3	1.461*** (0.0671)	1.180*** (0.0349)	0.325*** (0.0269)	0.295*** (0.0215)	0.414*** (0.0183)	0.663*** (0.0175)	0.680*** (0.022)	1.324*** (0.0531)
NUTS_2	0.168*** (0.0632)	0.200*** (0.032)	0,00793 (0.026)	0.0620*** (0.0208)	-0.0462*** (0.0165)	0.130*** (0.0164)	0.0665*** (0.0204)	-0,0177 (0.0489)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin Time-Varying F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination Time-Varying F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	195.026	195.026	195.026	195.026	195.026	195.026	195.026	195.026
R-squared	0,89	0,876	0,119	0,105	0,196	0,504	0,415	0,83

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 10: *Fixed effects estimation with Time.*

(a) *Total Value of Trade*

	Dist (0 Km - 150 Km)	*Dist (151 Km - 202 Km)*	*Dist (202 Km - 248 Km)*	*Dist (248 Km - 351 Km)*	*Dist (351 Km - 620 Km)*	*Dist (More than 620 Km)*	*General Specification*
Variables	**Total Value**	**Total Value**	**Total Value**	**Total Value**	**Total Value**	**Total Value**	**Total Value**
Distance	-0.0189*** (0.000909)	-0.00837*** (0.00234)	0.00422 (0.00276)	-0.00283** (0.00111)	-0.000932*** (0.000301)	-0.00120*** (0.000204)	-0.00537*** (0.000241)
Dist. Square	---	---	---	---	---	---	2.88e-06*** (1.76e-07)
Contiguity	0.955*** (0.0730)	---	---	---	---	---	1.309*** (0.0684)
GDPi	2.59e-08*** (6.88e-09)	3.55e-08*** (6.29e-09)	1.84e-08*** (5.24e-09)	2.27e-08*** (2.72e-09)	3.45e-08*** (4.57e-09)	1.88e-08*** (3.00e-09)	2.15e-08*** (2.70e-09)
GDP$_j$	2.67e-08*** (6.69e-09)	2.67e-08*** (5.64e-09)	4.18e-08*** (3.50e-09)	4.19e-08*** (7.30e-09)	3.28e-08*** (7.30e-09)	2.47e-08*** (3.72e-09)	2.61e-08*** (2.32e-09)
NUTS_5	2.782*** (0.0992)	---	---	---	---	---	3.178*** (0.0814)
NUTS_3	0.800*** (0.0868)	0.340 (0.248)	0.386 (0.364)	---	---	---	1.133*** (0.0645)
NUTS_2	0.388*** (0.0970)	0.175 (0.123)	0.358* (0.183)	0.258* (0.140)	0.101 (0.194)	---	-0.00289 (0.0619)
Observations	15,989	2,600	1,790	3,898	8,158	6,415	39,045
R-squared	0.960	0.942	0.946	0.852	0.779	0.688	0.941
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(b) *Extensive Margin*

	Dist (0 Km - 150 Km)	*Dist (151 Km - 202 Km)*	*Dist (202 Km - 318 Km)*	*Dist (318 Km - 439 Km)*	*Dist (439 Km - 761Km)*	*Dist (More than 439 Km)*	*General Specification*
Variables	**Shipments**	**Shipments**	**Shipments**	**Shipments**	**Shipments**	**Shipments**	**Shipments**
Distance	-0.0169*** (0.000512)	-0.0111*** (0.00146)	-0.00496*** (0.000653)	-0.00452*** (0.000605)	-0.00123*** (0.000175)	-0.000597*** (0.000122)	-0.00627*** (0.000173)
Dist. Square	---	---	---	---	---	---	3.28e-06*** (1.44e-07)
Contiguity	0.828*** (0.0373)	---	---	---	---	---	1.135*** (0.0350)
GDPi	2.82e-08*** (7.04e-09)	2.20e-08*** (6.46e-09)	1.92e-08*** (2.60e-09)	2.84e-08*** (5.01e-09)	2.36e-08*** (2.89e-09)	1.01e-08*** (3.80e-09)	2.49e-08*** (3.96e-09)
GDP$_j$	2.96e-08*** (7.52e-09)	3.34e-08*** (4.95e-09)	3.44e-08*** (2.66e-09)	3.89e-08*** (3.52e-09)	2.09e-08*** (2.97e-09)	7.74e-09*** (2.26e-09)	2.72e-08*** (5.27e-09)
NUTS_5	2.458*** (0.0532)	---	---	---	---	---	2.819*** (0.0475)
NUTS_3	0.669*** (0.0472)	0.356*** (0.131)	0.746*** (0.159)	---	---	---	0.970*** (0.0389)
NUTS_2	0.306*** (0.0581)	0.197** (0.0780)	0.203*** (0.0661)	0.230* (0.120)	---	---	0.0371 (0.0390)
Observations	15,989	2,600	4,509	4,309	8,059	3,498	39,045
R-squared	0.963	0.887	0.815	0.804	0.680	0.603	0.948
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(c) *Intensive Margin*

	Dist (0 Km - 150 Km)	*Dist (151 Km - 202 Km)*	*Dist (202 Km - 248 Km)*	*Dist (248 Km - 351 Km)*	*Dist (351 Km - 620 Km)*	*Dist (More than 620 Km)*	*General Specification*
Variables	**Average Value**	**Average Value**	**Average Value**	**Average Value**	**Average Value**	**Average Value**	**Average Value**
Distance	-0.00399*** (0.000397)	0.000639 (0.00159)	0.00441* (0.00239)	-0.000750 (0.000706)	-0.000258 (0.000181)	-0.000725*** (0.000104)	-0.000298*** (9.78e-05)
Dist. Square	---	---	---	---	---	---	-3.61e-08 (7.69e-08)
Contiguity	0.347*** (0.0338)	---	---	---	---	---	0.467*** (0.0312)
GDPi	6.98e-09** (3.47e-09)	2.25e-08*** (4.18e-09)	5.05e-09 (3.42e-09)	1.35e-08*** (3.26e-09)	1.35e-08*** (2.21e-09)	8.21e-09*** (2.14e-09)	9.05e-09*** (1.36e-09)
GDP$_j$	6.42e-09* (3.50e-09)	7.13e-09* (4.26e-09)	1.52e-08*** (4.22e-09)	2.15e-08*** (7.52e-09)	1.50e-08*** (3.57e-09)	1.06e-08*** (2.04e-09)	1.05e-08*** (1.50e-09)
NUTS_5	0.859*** (0.0578)	---	---	---	---	---	0.979*** (0.0609)
NUTS_3	0.264*** (0.0397)	0.395** (0.178)	-0.0170 (0.295)	---	---	---	0.313*** (0.0301)
NUTS_2	0.0946** (0.0414)	0.140* (0.0800)	0.353*** (0.106)	0.0723 (0.103)	-0.00280 (0.146)	---	0.0216 (0.0275)
Observations	15,989	2,600	1,790	3,898	8,158	6,415	39,045
R-squared	0.420	0.671	0.791	0.568	0.359	0.419	0.291
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 11: *Fixed effects estimation by Distance Breakpoints.*

(a) *Total Value of Trade*

Variables	Time (0 min - 103 min)	Time (103 min - 139 min)	Time (139 min - 178 min)	Time (178 min - 240 min)	Time (240 min - 509 min)	Time (More than 509 min)	General Specification
	Total Value	Total Value	Total Value	Total Value	Total Value	Total Value	Total Value
Time	-0.0243*** (0.00129)	-0.00429 (0.00335)	0.000489 (0.00356)	-0.00830*** (0.00214)	-0.00235*** (0.000415)	-0.000590*** (0.000115)	-0.00600*** (0.000290)
Time Square	---	---	---	---	---	---	3.16e-06*** (1.74e-07)
Contiguity	0.913*** (0.0751)	---	---	---	---	---	1.330*** (0.0692)
GDP_i	2.23e-08*** (8.48e-09)	3.21e-08*** (7.77e-09)	2.82e-08*** (4.74e-09)	1.70e-08** (6.64e-09)	2.87e-08*** (3.68e-09)	2.97e-08*** (4.73e-09)	2.08e-08*** (2.68e-09)
GDP_j	2.94e-08*** (8.45e-09)	1.70e-08** (8.54e-09)	3.18e-08*** (3.21e-09)	3.19e-08*** (4.79e-09)	3.41e-08*** (6.83e-09)	2.47e-08*** (5.49e-09)	2.54e-08*** (2.34e-09)
NUTS_5	2.801*** (0.107)	---	---	---	---	---	3.218*** (0.0837)
NUTS_3	0.834*** (0.0946)	0.613*** (0.234)	1.131** (0.445)	---	---	---	1.221*** (0.0642)
NUTS_2	0.393*** (0.109)	0.196* (0.113)	-0.177 (0.138)	-0.000873 (0.176)	-0.00476 (0.183)	---	0.0155 (0.0657)
Observations	15,064	2,948	2,485	3,001	10,016	5,531	39,045
R-squared	0.958	0.940	0.938	0.733	0.766	0.764	0.941
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(b) *Extensive Margin*

Variables	Time (0 min - 103 min)	Time (103 min - 138 min)	Time (138 min - 180min)	Time (180 min - 423 min)	Time (423 min - 528 min)	Time (More than 528 min)	General Specification
	Shipments	Shipments	Shipments	Shipments	Shipments	Shipments	Shipments
Time	-0.0227*** (0.000746)	-0.00824*** (0.00222)	-0.0405*** (0.000637)	-0.00332*** (0.000205)	-0.00339*** (0.00109)	-0.000535*** (8.37e-05)	-0.00757*** (0.000195)
Time Square	---	---	---	---	---	---	4.03e-06*** (1.17e-07)
Contiguity	0.775*** (0.0381)	---	---	---	---	---	1.147*** (0.0350)
GDP_i	3.05e-08*** (8.60e-09)	1.79e-08* (9.42e-09)	3.06e-08*** (4.18e-09)	2.20e-08*** (3.00e-09)	2.75e-08*** (3.75e-09)	1.70e-08*** (2.99e-09)	2.40e-08*** (4.05e-09)
GDP_j	2.71e-08*** (9.04e-09)	8.28e-09 (6.86e-09)	3.41e-08*** (5.11e-09)	3.11e-08*** (4.57e-09)	2.54e-08*** (3.64e-09)	1.83e-08*** (2.45e-09)	2.66e-08*** (5.32e-09)
NUTS_5	2.435*** (0.0558)	---	---	---	---	---	2.782*** (0.0484)
NUTS_3	0.676*** (0.0497)	0.278** (0.130)	-1.870*** (0.0538)	---	---	---	0.998*** (0.0385)
NUTS_2	0.317*** (0.0636)	0.220*** (0.0735)	-1.557*** (0.0688)	0.117 (0.0715)	---	---	0.0165 (0.0387)
Observations	15,064	2,903	18,880	9,376	4,124	4,920	39,045
R-squared	0.962	0.877	0.961	0.673	0.795	0.606	0.950
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

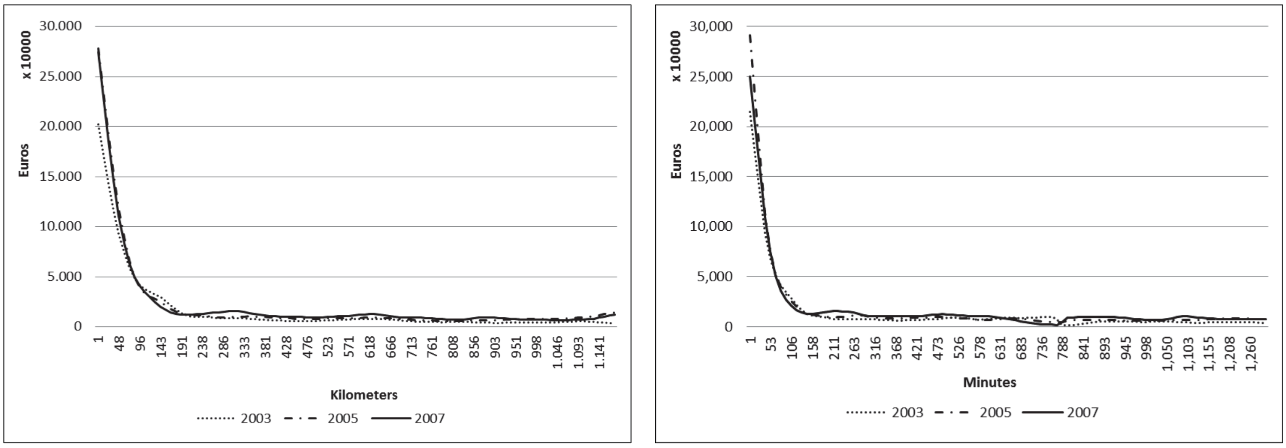
(c) *Intensive Margin*

Variables	Time (0 min - 103 min)	Time (103 min - 139 min)	Time (139 min - 178 min)	Time (178 min - 240 min)	Time (240 min - 509 min)	Time (More than 509 min)	General Specification
	Average Value	Average Value	Average Value	Average Value	Average Value	Average Value	Average Value
Time	-0.00698*** (0.000578)	-0.00150 (0.00239)	0.00243 (0.00270)	-0.00104 (0.00148)	-0.000166 (0.000178)	-0.000328*** (5.90e-05)	-0.000541*** (9.38e-05)
Time Square	---	---	---	---	---	---	2.40e-07*** (5.46e-08)
Contiguity	0.289*** (0.0346)	---	---	---	---	---	0.463*** (0.0311)
GDP_i	4.73e-09 (4.24e-09)	1.87e-08*** (4.44e-09)	1.56e-08*** (3.22e-09)	4.90e-09 (5.39e-09)	1.33e-08*** (1.96e-09)	8.79e-09*** (2.23e-09)	9.11e-09*** (1.36e-09)
GDP_j	8.96e-09** (4.31e-09)	1.64e-08*** (5.11e-09)	8.10e-09** (3.23e-09)	1.29e-08** (5.45e-09)	1.58e-08*** (3.79e-09)	1.14e-08*** (2.89e-09)	1.05e-08*** (1.50e-09)
NUTS_5	0.783*** (0.0596)	---	---	---	---	---	0.978*** (0.0613)
NUTS_3	0.200*** (0.0417)	0.318** (0.133)	0.791** (0.399)	---	---	---	0.317*** (0.0291)
NUTS_2	0.0927** (0.0466)	0.222*** (0.0780)	0.187* (0.0965)	0.0508 (0.114)	-0.0811 (0.113)	---	0.0341 (0.0276)
Observations	15,064	2,948	2,485	3,001	10,016	5,531	39,045
R-squared	0.423	0.652	0.742	0.606	0.354	0.417	0.291
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors. Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

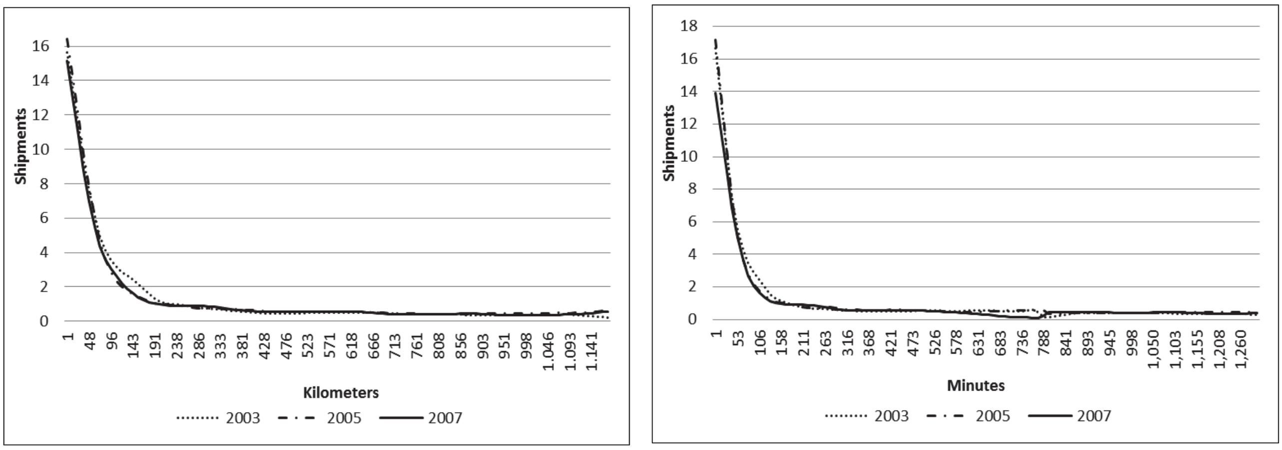
Table 12: *Fixed effects estimation by Time Breakpoints.*

B. Kernel regressions: First level decomposition.



(c) *Distance* (d) *Travel Time*

Figure 9: *Kernel regressions: Total value of trade on distance and time.*



(e) *Distance* (f) *Travel Time*

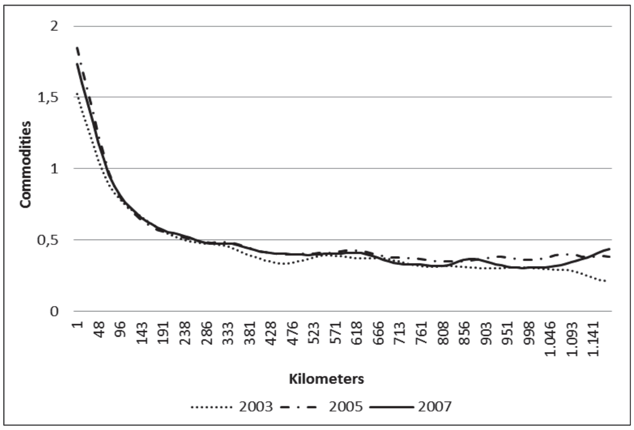
Figure 10: *Kernel regressions: Number of shipments (extensive margin) on distance and time.*

(g) *Distance*

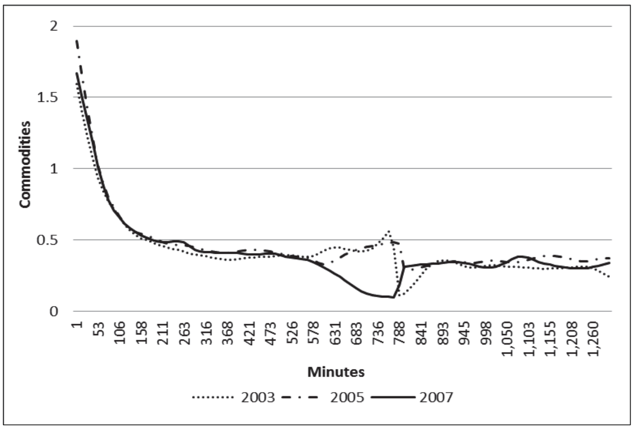
(h) *Travel Time*

Figure 11: *Kernel regressions: Average shipment value (intensive margin) on distance and time.*

C. Kernel regressions: Second level decomposition

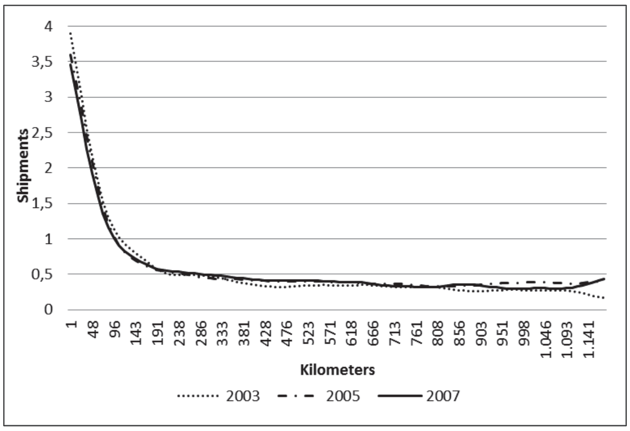


(i) *Distance*

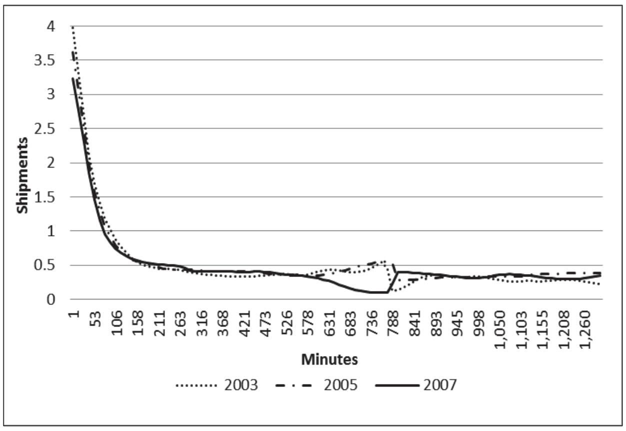


(j) *Travel Time*

Figure 12: *Kernel regressions: Number of commodities on distance and time.*

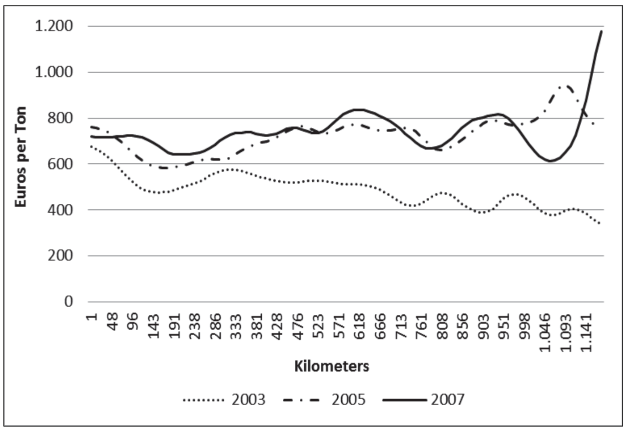


(k) *Distance*



(l) *Travel Time*

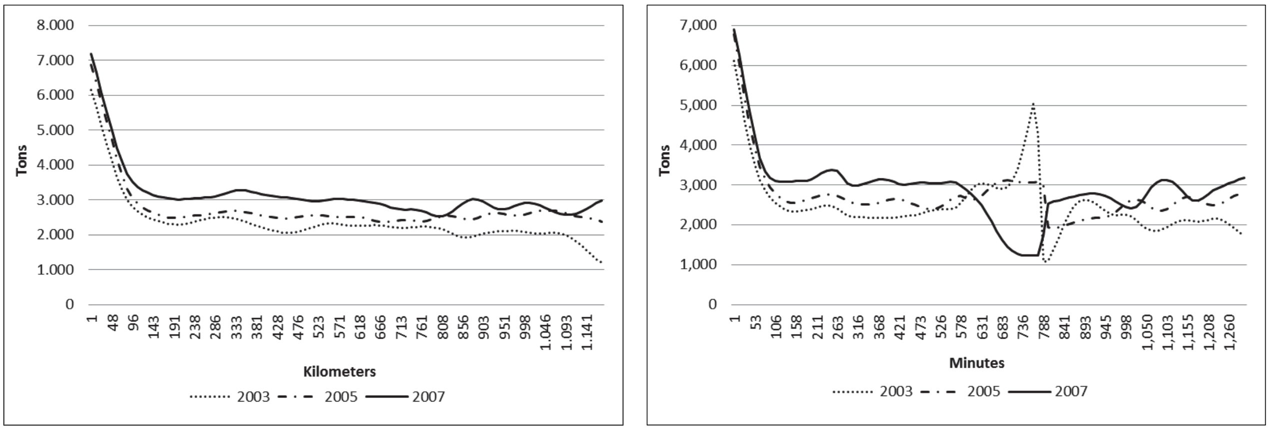
Figure 13: *Kernel regressions: Number shipments per commodity (frequency) on distance and time.*



(m) *Distance*

(n) *Travel Time*

Figure 14: *Kernel regressions: Average price on distance and time.*



(o) *Distance* (p) *Travel Time*

Figure 15: *Kernel regressions: Average quantity (tons) on distance and time.*

D. Kernel regressions: Madrid and Barcelona trade flows.

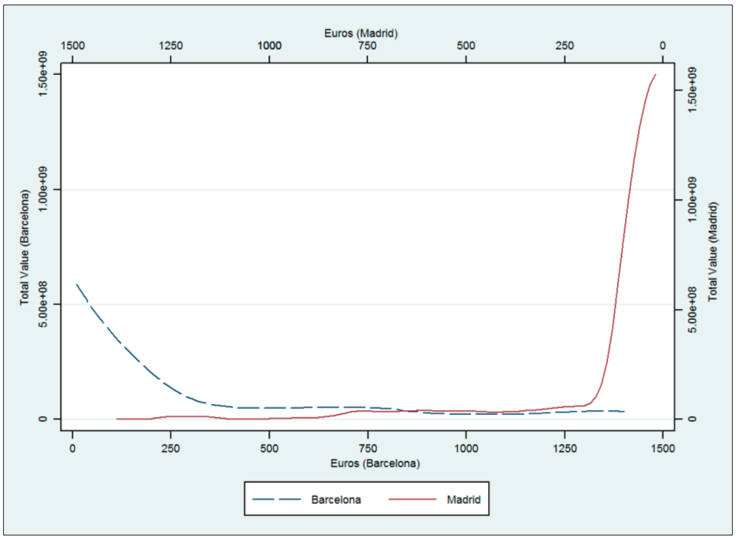


Figure 16: *Trade areas competition for Madrid and Barcelona.*