European
Commission

JRC TECHNICAL REPORT

A trade hierarchy of cities based on transport cost thresholds

JRC Working Papers on
Territorial Modelling and Analysis
No 02/2019

Authors:
Díaz-Lanchas, J.,
Llano, C.,
Zofío, J.L.

2019

Joint
Research
Centre

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information
Name: Simone Salotti
Address: Edificio Expo, C/Inca Garcilaso 3, 41092 Sevilla (Spain)
Email: simone.salotti@ec.europa.eu
Tel.: +34 954488250

EU Science Hub
https://ec.europa.eu/jrc

JRC115750

Seville: European Commission, 2019

© European Union, 2019

The reuse policy of the European Commission is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed provided appropriate credit is given and any changes are indicated. For any use or reproduction of photos or other material that is not owned by the EU, permission must be sought directly from the copyright holders.

All content © European Union, 2019 (unless otherwise specified)

How to cite this report: Diaz-Lanchas, J., Llano, C., and Zofio, J.L. (2019). A trade hierarchy of cities based on transport cost thresholds. JRC Working Papers on Territorial Modelling and Analysis No. 02/2019, European Commission, Seville, JRC115750.

The **JRC Working Papers on Territorial Modelling and Analysis** are published under the supervision of Simone Salotti and Andrea Conte of JRC Seville, European Commission. This series mainly addresses the economic analysis related to the regional and territorial policies carried out in the European Union. The Working Papers of the series are mainly targeted to policy analysts and to the academic community and are to be considered as early-stage scientific papers containing relevant policy implications. They are meant to communicate to a broad audience preliminary research findings and to generate a debate and attract feedback for further improvements.

A Trade Hierarchy of Cities based on Transport Cost Thresholds

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

[a] *Joint Research Centre - European Commission*
[b] *Universidad Autónoma de Madrid and Lawrence. R. Klein Institute*
[c] *Universidad Autónoma de Madrid and Erasmus Research Institute of Management.*

Abstract

Empirical evidence has been lacking to explain trade agglomerations over short distances. Starting with a novel micro-database of road freight shipments within Spain for the period 2003-2007, we break down city (municipal) trade flows into the extensive and intensive margins and assess trade frictions and trade concentration relying on a unique generalized transport cost measure and three internal borders, NUTS-5 (municipal), NUTS-3 (provincial) and NUTS-2 (regional). We discover a stark accumulation of trade flows up to a transport cost value of €189 (approx. 170km) and conclude that this high density is due not to administrative borders effects but to significant changes in the trade-to-transport costs relationship. To support this hypothesis, we propose and conduct an endogenous Chow test to identify significant thresholds at which trade flows change structurally with distance. These breakpoints allow us to split the sample when controlling for internal borders, and to define trade market areas corresponding to specific transport costs values that consistently reveal an urban hierarchy of cities. The results provide clear evidence to corroborate the predictions of central place theory.

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

JEL Classification Numbers: F14, F15, O18

---

1. Introduction

In absence of reliable data, intra-national trade patterns are in general neither well understood, nor well characterized. The lack of highly granular micro-data on interregional trade flows deters empirical analyses on why trade-agglomerations occur in some places, and not in others within a country.[1] Since the first empirical and seminal studies (Wolf, 1997, 2000; Hillberry and Hummels, 2003) up to more recent works (Crafts and Klein., 2014; Borraz et al., 2016; Martínez-San Román, et al., 2017), there is growing literature pointing out internal home bias effect as, inter alia, one of the determinants of intra-national trade agglomerations.

---

[*] Corresponding author: Jorge.DIAZ-LANCHAS@ec.europa.eu.

[1] Although the lack of official databases on intra-national trade flows discourages empirical trade analysis, recent and prominent works provide estimates that apply context-specific techniques to survey data (Wolf, 2000; Hillberry, 2002a; Hillberry and Hummels, 2008), use other official data sources at the national level (Poncet, 2005; Llano et al., 2010; Yilmazkuday, 2011, 2012) or rely on input-output databases (Thissen, et al., 2014).

The causes and explanations for this intra-country effect are multiple. Whereas some studies focus on historical barriers to trade (Nitsch and Wolf, 2013; Wrona, 2018), hub-spoke structures (Gallego, et al., 2015), transport-mode competitions within a country (Llano et al., 2017), or even the impediments raised by business networks (Garmendia, et al., 2012) and political boundaries (Borraz et al., 2016), others emphasize unexpected increasing in the internal border effect over the last decades (Crafts and Klein., 2014; Martínez-San Román, et al., 2017). Not only these empirical factors but also prominent theoretical set-ups have been cited to ascribe the internal home-bias effect either to the fragmentation of global production chains in areas near international borders (Yi, 2010), or to spatial frictions between cities (Behrens et al., 2017).

Within this framework, the study by Hillberry and Hummels (2008) sheds light on the appearance of an internal home bias at the municipal level once accurate interregional trade flows and very precise measures of internal distance are used. Even if this trade concentration holds over short distances in their study, they conclude that the internal home bias is a *reductio ad absurdum* of the one observed at the international level. Starting with this point we argue that the internal home bias effect not only is real, but results from breakpoint-changes in the trade-to-transport costs relationship that shapes a series of trade market areas around a country's largest trading-cities. More specifically, we find changes in trade-to-transport costs patterns that map successive market areas monotonically with distance–areas that grow larger in tandem with the size of the location. The result is a hierarchy of cities as predicted by central place theory (Parr, 2002; Tabuchi and Thisse, 2011; Hsu, 2012; Mulligan et al., 2012; Hsu et al., 2014).

To arrive at these conclusions we have relied on two unique and very detailed databases: *i*) a panel dataset (2003-2007) on individual trade flows for Spanish municipalities; and *ii*) an additional bilateral panel database with the precise and realistic transport costs associated with these trade flows. With these databases we propose and adopt a new methodological approach within the trade literature to determine breakpoints in spatial trade trends caused by the changing nature of transport costs associated with individual shipments. The new method resorts to an endogenous econometric (Chow) test for structural changes, first introduced by Berthelemy and Varoudakis (1996). We apply this test to assess whether trade flows (structurally) change beyond a certain distance-breakpoint. This information allows us to split internal distances by these thresholds and introduce them into a standard gravity setup. This new definition of internal distances is the basis, on one hand, to define cities' market areas and, on the other, to accurately measure the internal home bias at different spatial levels, for either the overall value of trade or its margins. Whereas we interpret the former under the insights of central place theory and the consequent hierarchy of cities, we show that the latter leads to an overestimation of the ("illusory") internal border effect if transport costs are not correctly controlled for. Indeed, we argue that the arbitrary use of administrative boundaries as a spatial limitation to collect data and account for border effects becomes an *aggregation artifact* (Hillberry, 2002a; Llano-Verduras et al., 2011) that leads to misleading conclusions about the impediments to intra-national trade flows.

This paper contributes to and links the literatures on intra-national trade flows and their associated system of city mapping. We argue and illustrate that it is not internal administrative borders that effectively hamper trade, but the change in the relationship between the trade and transport costs what generates the illusory border effects achieved within a country. This is consistent with the removal of intra-national administrative impediments to trade, through single-market agreements. We also contribute to the empirical study of trade-market areas as intra-national economic boundaries depending on either their geographical reach (Löffler, 1998) or their shape throughout a hierarchy of cities (Hsu, 2012; Hsu et al., 2014), and provide empirical evidence for the prominent and theoretical literature on trading-cities (Anas and Xiong, 2003; Cavailhes et al., 2007; Behrens et al., 2017; Mori and Wrona, 2018). Finally, we develop further methodological approaches for highly granular gravity-trade models (Head and Mayer, 2014) and their crucial debate on the treatment of non-linear transport costs (Eaton and Kortum, 2002; Abbate, et al., 2012; Gallego and Llano, 2014).

Now that we have reviewed the paper's contributions, we outline the sequential research strategy. The first step was to compile a novel database of road freight shipments consisting of micro-data for individual deliveries between Spanish municipalities. This statistical information, drawn from the Spanish Road Freight Transportation Survey (RFTS), allows us to calculate the bilateral value of trade and decompose it into the extensive and the intensive margins. Next, for each bilateral flow we computed the actual monetary transport cost, or generalized transport cost (GTC), which corresponds to the minimum economic cost between any origin and destination. This is the sum of distance-related costs (e.g., fuel, toll, tires) and time-related costs (e.g., salaries, insurance, taxes) for each route between any pair of Spanish municipalities. To calculate these we used the programming techniques available in Geographic Information Systems (Arc/GIS), which can optimize for the lowest -cost routes using the existing and digitalised road networks for 2003-2007. In contrast to all previous studies, which use the standard and non-monetary transport cost proxies of distance or travel time, we introduce a *real euro* time-varying measure of the spatial frictions affecting trade.

Subsequently, we analyzed transport costs and border impediments for the three intra-national *administrative* borders in Spain–i.e, NUTS-5 (municipal), NUTS-3 (provincial) and NUTS-2 (regional)– using the Poisson pseudo-maximum likelihood (PPML) estimator to control for pervasive zero-valued trade flows and heteroskedasticity problems within the gravity model (Santos Silva and Tenreyro, 2006, 2010, 2011). The results show that municipal borders have a stronger impact on trade flows and extensive and intensive trade margins than the results reported by, among others, Hillberry and Hummels (2008). Provinces and regions, meanwhile, have a much weaker impact than the literature suggests (Requena and Llano, 2010). This spatial trade pattern is explained by the over-concentration of trade flows over short distances (measured in euros) that only a PPML estimator can capture. In consequence, as we show, the internal effect of administrative borders tends to be underestimated for short distances, and overstressed for long ones.

To explain these results further we transpose to the trade literature the method developed by Berthelemy and Varoudakis (1996) for endogenous economic growth models. These authors propose an endogenous Chow test to detect specific *breakpoints* shaping different structural models in their data. We adopt this test to determine differences in the trade-to-transport costs relationship so that we can then divide our transport costs measure into sub-samples (distance-intervals). We hope this first use of the structural-break methodology in the trade literature will illustrate its potential to define *natural* market areas. Once we had determined the existence of a series of statistically significant breakpoints, we split our PPML regressions by these new thresholds. This stands in contrast with, for example, Eaton and Kortum (2002) and Anderson and Yotov (2012), who segment trade flows by arbitrary distances. Thanks to these statistically defined transport-cost thresholds, we conclude both that the internal border effect is neither unique, nor has a single non-linear quadratic impact on trade, as the literature emphasizes. Indeed, we argue that it is precisely the existence of such thresholds that truly raise internal impediments to trade, which eventually spill over several and consecutive administrative boundaries.

Lastly, since the density of trade flows is geographically localized in short-distance thresholds, we show that the resulting agglomeration economies form a series of trade-market areas around increasingly large cities, giving rise to a urban hierarchy whose mapping coincides with the postulates of central place theory. Thus the so-called internal border effect is the consequence of high-order cities that act as supply centers for either their metropolitan areas or surrounding cities of lower-rank, a relatively unexplored relationship in the empirical literature on trade and urban areas. These empirical regularities are promising for future research linking trade and central place theory.

We now summarize the structure of the paper. Section 2 explains the databases and the decomposition of trade flows. Section 3 performs the econometric analysis. Section 4 discusses the structural-breakpoint methodology and its application to the gravity model of trade. Section 5 interprets our results in line with central place theory. The last section draws the main conclusions.

2. Trade flows, extensive and intensive margins, and transport costs.

2.1. Trade value data and the extensive and intensive margins.

To build our value trade data we start with a micro-database on physical shipments: the Road Freight Transportation Survey (RFTS), compiled by the transport division of the Spanish Ministerio de Fomento for the period 2003-2007. This dataset is based on a randomly surveyed sample of freight companies and independent truckers with vehicles weighing more than 3.5 tons, which together account for almost 85% of all internal freight flows. It includes such vehicle and shipment characteristics as the type of product shipped, the number of tons carried by the truck, the number of shipments and the distance travelled between any origin i and destination j, registered at the NUTS-5 municipal level.[2] However, the survey does not provide information on the value of traded goods. We therefore need product prices (in €/ton) that we can multiply by tons carried, to yield an estimated total value. Because of statistical confidentiality, these prices are not available at the municipal level (NUTS-5) in any of the official databases. To obtain a representative price-vector for the period, then, we rely on an alternative interregional trade flows database compiled by the C-intereg project (Llano et al., 2010).[3]

Relying on the RFTS micro-database we calculate the total value of trade and subsequently break it down into the extensive and the intensive margins. According to Hillberry and Hummels (2008) proposal, the total value of trade between each origin-destination (T_{ijt}) at time t can be decomposed as follows (dropping subscript t to avoid notational clutter):

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

where N_{ij} represents the total number of shipments (i.e., the extensive margin) and $\overline{PQ}_{ij}$ is the average value per shipment (i.e., intensive margin) composed of the mean of values of prices ($\overline{P}_{ij}$) and tons ($\overline{Q}_{ij}$).[4]

---

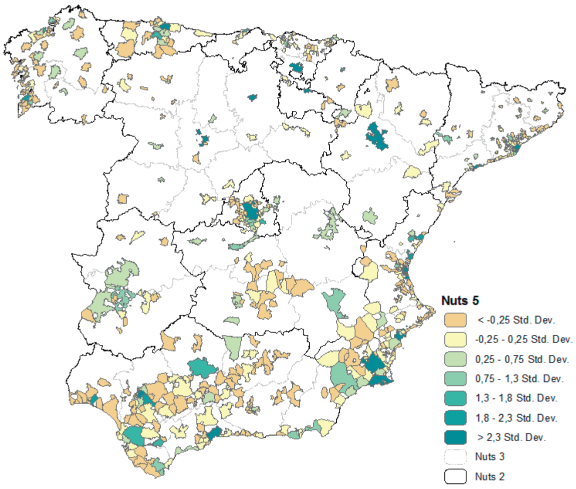
[2] The RFTS differs in several ways from the American Commodity Flow Survey undertaken jointly by the Census Bureau and the Bureau of Transportation Statistics (DOT). It is relevant for this study that the surveyed statistical units in the latter are production establishments (wholesalers and retailers), while the RTFS surveys road freight companies providing the transport service.

[3] However, these prices are calculated at the provincial level (NUTS-3). We must therefore assume that prices at the municipal and provincial levels are equal. This assumption implies that the pricing rules at municipal level–e.g., mark-ups over costs–are similar to those observed at the provincial level.

[4] In Appendix A.1, online, we explore our series of results for a second-level decomposition of the extensive and intensive margins, so as to draw further conclusions on specific trade patterns.

Figure 1 maps the standard deviation of total value of trade for the whole period and for the 633 municipalities with more than 10.000 inhabitants. The largest trade flows occur in areas with the highest levels of economic activity (Madrid, the Mediterranean coast and the Basque Country), whereas less populated areas (south-west and north-west) register trade only around largest cities. This spatial pattern being consistent with previous evidence on the concentration of Spanish international exporters (Ramos and Moral-Benito, 2017).

Figure 1: Total value of trade for the 633 municipalities. Average values for 2003-2007.



Source: Own elaboration from the RFTS data.

2.2. Generalized transport cost (GTC). Advantages over distance and time proxies.

Another novel aspect of this study is the use of a monetary measure of transport cost that improves those normally used as rough transport costs approximations, mainly geographical distance and travel time. This variable stands for a generalized transport cost (GTC) definition corresponding to the least-cost itinerary between an origin and a destination, as discussed in Combes et al. (2005) and Zofío et al. (2014). The GTC is calculated with GIS software (Arc/GIS) using digitalized road networks for 2003-2007. GTCs differentiate economic costs related to both distance and time. The distance-economic cost (in 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 repair operating costs. On the other side, the time-economic cost (euros per minute) includes the following variables: labor costs (gross salaries), financial costs associated with the amortization, insurance costs, taxes, truck financing, and indirect costs associated with other operating expenses including administration and commercial costs.[5] Unlike distance and travel time, the GTC accounts for changes in transport operating costs, particularly those related to the transport service market (e.g., fuel costs) and regulatory conditions (e.g., mostly wages in the labor market).

---

[5] The optimal minimum cost itinerary among the set of possible itineraries is defined as follows: $GTC_{ij} = min(Distance\ Economic\ Cost_{ij} + Time\ Economic\ Cost_{ij})$.

Table 1: Transport costs variation. Average values.

Transport costs	**2003-2007**	**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%

Source: Own elaboration from the GTC database.

Table 2: Variation effect of transport costs on trade flows.

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

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

Throughout the text we argue that a time-variant measure of transport costs, such as the GTC, is more suitable for trade flows analyses than distance and travel time because it is the only measure that captures simultaneous improvements in road transport infrastructure as well as changing regulations in time. Table 1 shows the average value of each transport cost for the period 2003-2007 and the individual years 2003, 2005 and 2007, plus their growth rates. As shown, the GTC is the measure with the highest variation. Distance remains mainly unaffected by road improvements over the years. Physical distance can constitute a good proxy for transport costs in cross–sectional studies. Because of its lack of variability, however, it is certainly inadequate in a panel data setting, where transport costs are expected to vary significantly. Correspondingly, whereas the travel time proxy for transports costs captures the improvements in road infrastructure, it cannot account for changes in operating costs.

Additionally, to reinforce the need for a GTC definition instead of its distance and time proxies, we regress the growth rate of trade flows and their margins (ΔT_{ijt}) on the growth rates of 'transport costs' ($\Delta Cost_{ijt}$)—representing either GTC, distance or travel-time—and including also a year variant and origin-destination invariant fixed effects:

$$\Delta T_{ijt} = \beta_0 + \beta_1 \Delta Cost_{ijt} + year + \eta_i + \eta_j + \epsilon_{ijt} \tag{2}$$

Table 2 shows the results. As expected, distance and travel time do not have significant effects neither on the growth rates of trade flows nor on their margins. By contrast, the GTC has a negative and significant impact on the three dependent variables. That is, reductions in the GTC–such as those related to truck efficiency, fuel saving or even reductions in salaries–leads to an increase in trade flows.

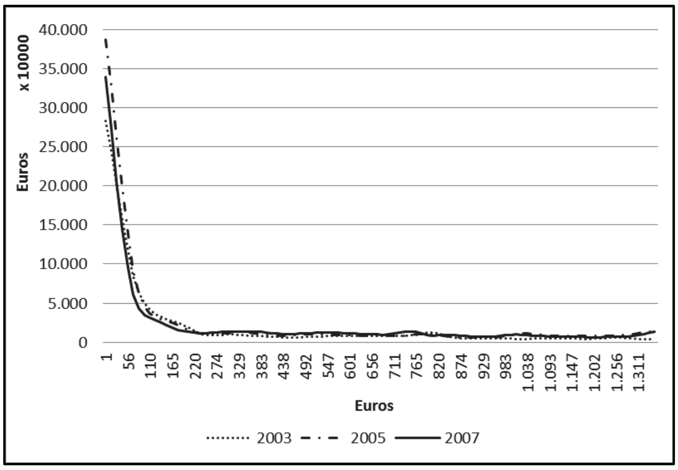
2.3. Trade densities on transportation costs.

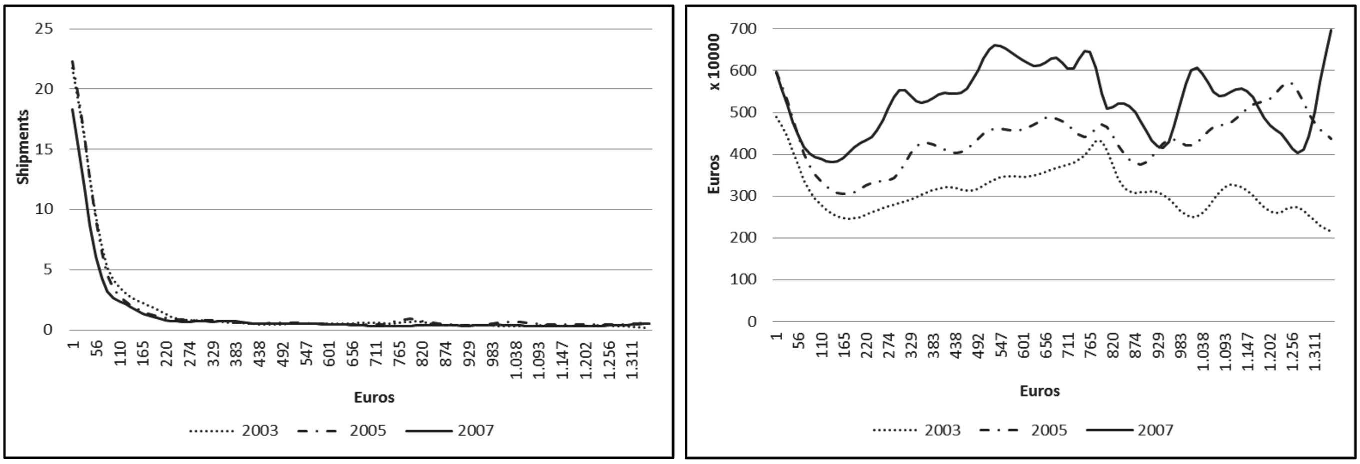
We now resort to non-parametric estimation (kernel regressions) to study spatial trade dynamics using the GTC. Figures 2a, 2b, and 2c present the kernel regressions for each component in Eq.(1)

respectively. [6] Total value of trade—marginally lower in 2003 than in 2007 as a result of the starting recession—falls sharply in density as transport costs increase (2a). This pattern is driven by the extensive margin (2b) where the number of shipments drops rapidly for all years up to a GTC threshold value at which their non-linear relation is observed. The intensive margin (2c), on the other hand, shows an upward trend even with increasing transport costs. This increasing trend is due to the price-component—see Figure A.2 in the online Appendix, whereas tons are highly concentrated in short distances, reflecting the extensive margin behavior and, mostly, the accumulation of shipments around major metropolitan areas (Madrid, Barcelona and Valencia).

Figure 2: Kernel regressions.

(a) Total Value of Trade




(b) Extensive Margin

(c) Intensive Margin

---

[6] We use the Gaussian kernel estimator, with $n = 100$ points, and allow it to calculate and employ the optimal bandwidth.

3. Econometric specification: Trade frictions.

To structure the analysis, in this section we first explore the effects of trade frictions on the total value of trade and its margins by defining a standard gravity model in conjunction with a Poisson pseudo-maximum likelihood (PPML) estimation.[7] For this, we propose a set of regressions relying on trade decompositions from Eq.(1) and using Eq.(3) for the PPML estimator. We include trade flows and their margins in levels (X_{ijt}) and transport costs in a quadratic way (GTC_sq_{ijt}) to capture likely non-linear effects of transport costs (McCann, 2005). To accurately isolate the effect of municipal border the contiguity variable is calculated as a first-order queen contiguity that takes the value one when i and j share a border but also when flows take place within the same municipality (Hillberry, 2002b). For internal administrative boundaries, we consider three dummy variables following Requena and Llano (2010). Consequently, NUTS-5 (municipalities) is set to one when flows are performed within the same municipality and to zero otherwise, whereas the NUTS-3 (provinces) border takes the value one when flows are in the same province but not in the same municipality, and the NUTS-2 (regions) border captures flows between two municipalities that are located in different provinces but in the same NUTS-2 region. Finally, time, origin and destination fixed effects are included (Baldwin and Taglioni, 2006). Thus, the final specification to estimate corresponds to:

$$\begin{gathered}\mathrm{X}_{ijt} = \beta_0 + \beta_1 GTC_{ijt} + \beta_2 GTC\_sq_{ijt} + \beta_3 Contiguity + \\ \beta_4 NUTS\text{-}5 + \beta_5 NUTS\text{-}3 + \beta_6 NUTS\text{-}2 + year + \eta_i + \eta_j + \epsilon_{ijt.}\end{gathered} \tag{3}$$

Table 3 shows the results. GTC coefficients exhibit the expected negative sign in the linear component and signal increasing returns in transport (i.e., a positive sign in the quadratic term). For internal borders, trade within the same municipality (NUTS-5) is much greater than inter-municipal trade flows. Indeed, the extensive margin achieves a higher NUTS-5 coefficient than the intensive one, indicating that the extensive margin is what truly drives intra-municipal trade flows. Other administrative levels (NUTS-3 and NUTS-2) reduce their importance, especially regional flows. It turns out that border effects for relatively long-distance flows between regions are quite reduced or almost negligible; this reinforces the role of local markets and the extensive margin in trade agglomerations over short distances. Finally, these results (border coefficients and goodness of fit, R^2) are by far greater than—and in contrast with—those previously obtained by Hillberry and Hummels (2008). Given the number of properties satisfied by the PPML estimators over their OLS counterparts, we are confident that these results reliably explain and represent trade flows over very short distances. Against this estimation background, the use of granular measures for trade flows and transport costs reduces the impact of higher regional borders, showing the relevance of the municipal border. That is, these estimations confirm the “illusory effect” of regional borders once

---

[7] The PPML method accounts for the fact that the sample may include a large number of zeros, offering estimators that are more efficient than their OLS counterparts. This method identifies and eventually drops regressors that may cause the non-existence of (pseudo) maximum likelihood estimates. It therefore presents several advantages for trade data in light of the problems posed by the presence of numerous zeros, and in the presence of dummies–see also Head and Mayer (2014). In addition, the PPML can resolve the bias caused by heteroscedasticity, serial correlated error and multicollinearity, because of a high correlation among determinants in the gravity equation and the origin-destination dummies.

Table 3: PPML estimations with GTC and administrative borders.

Variables	Total Value of Trade (T_{ij})	Extensive Margin (N_{ij})	Intensive Margin ($\overline{PQ}_{ij}$)
GTC	-0.266***	-0.290***	-0.0423***
	(0.0322)	(0.0142)	(0.00537)
GTC Sq.	9.06e-11***	9.99e-11***	1.27e-11***
	(1.70e-11)	(7.40e-12)	(3.12e-12)
Contiguity	1.365***	1.126***	0.458***
	(0.0708)	(0.0330)	(0.0311)
NUTS-5	3.764***	3.236***	0.983***
	(0.101)	(0.0496)	(0.0594)
NUTS-3	1.717***	1.406***	0.320***
	(0.0840)	(0.0392)	(0.0255)
NUTS-2	0.346***	0.338***	0.0375
	(0.0777)	(0.0354)	(0.0259)
Year FE	Yes	Yes	Yes
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Observations	195,026	195,026	195,026
R^2	0.708	0.879	0.087

Clustered standard errors by *ij*. Std errors in parentheses
Significance level: ***p<0.01,**p<0.05, *p<0.1.

lower administrative limits are controlled for. Also, they indicate a non-disruption of the market for large administrative levels, in contrast to other studies for the Spanish (Requena and Llano, 2010; Garmendia, et al., 2012) and Chinese (Poncet, 2005) cases.

Moving forward, we also study trade dynamics in our panel database. We re-estimate Eq.(3) but resorting to a cross-section analysis for 2003 and 2007 independently. Focusing on the relevant results, Table 4 reports only the coefficients for administrative borders. All administrative levels reflect the same pattern for total trade and its margins; i.e., there exists a slowdown trend between 2003 and 2007.[8] Note that the NUTS-5 border in 2003 becomes even larger than in the panel data regressions, whereas the almost null (non-significant) impediments to trade of the NUTS-2 border are confirmed. These findings reflect the time-varying behavior of the internal home bias effect; that is, internal borders do not constantly deter the same amount of trade. Contrary to previous evidence for the US case (Martínez-San Román, et al., 2017), Spain experienced a decreasing internal border effect leading to a more trade integration at least during the economic growth of the 2003-2007 period.

---

[8] We further perform a set of mean tests to analyze whether each administrative border is statistically different between 2003 and 2007, rejecting the null hypotheses of equal coefficients.

Table 4: Cross-section regressions for 2003 and 2007.

GTC	Border	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

Clustered standard errors by *ij*. Std errors in parentheses
Significance level: ***p<0.01,**p<0.05, *p<0.1.

4. Structural trade patterns: GTC thresholds and the gravity model.

4.1. Breakpoints in the trade flows-to-transport cost relationship.

Previous regressions provide an estimate of the internal border effect qualified by the non-linear effect of transport costs and its customary quadratic second-order effect. Nevertheless, this specification does not consider the possibility of having alternative trade frictions once we control for structural changing patterns in trade flows. Visual inspection of the kernel regressions in Figure 2 reveals at least two different trade flows-transport cost structures around a critical value in the interval €110-€190, where their no-linear nature is clearly visible. This value implies that in very short distances trade flows are extremely sensitive to relatively low transport costs, particularly in their extensive margin, whereas they are almost independent ("becoming flat") for larger GTC. Moreover, from this GTC-range to larger values, further and multiple GTC thresholds may coexist at different levels. It is our understanding that we should test and control for these hypothetical GTC values. That is, we should assess whether trade flows (structurally) change once they reach one (or several) transport cost thresholds or *breakpoints*.

As remarked by Abbate, et al. (2012), previous works in the literature (Eaton and Kortum, 2002; Anderson and Yotov, 2012; Gallego and Llano, 2014) propose different distance intervals to determine the magnitude of trade barriers, although they define them arbitrarily, with a subjective criterion in terms of transport cost proxies. We propose to rely on statistical methods to identify these intervals. To determine these structural breakpoints we apply an endogenous Chow structural test for cross-sectional studies to check the stability of our parameters of interest.[9] We perform the test by adopting,

---

[9] This test works as follows. First we divide the sample into two sub-samples. Second, we carry out several iterations modifying the number of observations in each sub-sample, until we find a breakpoint in the relation between the dependent variable (trade flows) and our threshold variable (GTC). In this case the Chow test rejects the null hypothesis of the non-existence of a structural change in our model (at the 5% significance level). As argued by Diallo (2014), the advantage of this test is that it does not require us to provide in advance an exogenous point where we suspect a structural change, as in time-series models. Instead, it endogenously determines the exact (GTC) breakpoint at which the dependent variable changes significantly. This test has been recently made available in STATA by Diallo

first, a cross-section model that allows for the existence of a structural change. Eq.(3) represents the baseline model to develop such an analysis, with trade flows and GTCs as the dependent and threshold variables, respectively, to identify structural breakpoints. Also, we do not include the quadratic term of the GTCs, because in the methodology the non-linear relationships are captured by the transport cost thresholds. Once a GTC breakpoint is achieved, we take this corresponding GTC value and perform a second Chow test over trade flows above the GTC breakpoint previously obtained. Performing the test sequentially in this way, we can check for the existence of successive thresholds.

To summarize the information and rely on single thresholds for the whole sample, instead of performing the test for each year, we take the average of each variable in Eq.(3) for the whole period (2003-2007) and run the Chow test. Table 5 shows the breakpoints obtained for the total value of trade and the extensive margin. Unsurprisingly, the test fails to detect breakpoints for the intensive margin as its components do not drive reductions in trade as shown in the kernel regressions (Fig.2c). However, we confirm the existence of multiple GTC thresholds across the full spectrum of distances, providing strong evidence that trade flows are driven by the extensive margin, especially for short and medium GTCs (the first and second breakpoints correspond mainly to the total value of trade and its extensive margin). Moreover, trade flows are highly concentrated at low transport cost values (around €189 and €233), but over longer distances the difference between breakpoints becomes larger, indicating a decline in transport costs as an impediment to trade flows.

Table 5: GTCs thresholds for the value of trade and the extensive margin. Average values 2003-2007.

	GTC Break Points (in euros)				
Variable	Break Point 1	Break Point 2	Break Point 3	Break Point 4	Break Point 5
Total Value of Trade	189***	233***	285***	513*	706***
Extensive Margin	185***	246***	321***	582***	655***

Clustered standard errors by *ij*. Std errors in parentheses
Significance level: ***p<0.01,**p<0.05, *p<0.3.

Figures 2a and 2b show geographical evidence of the existence of these breakpoints by mapping the specific GTC breakpoints obtained for the two largest Spanish cities: Barcelona and Madrid.[10] The Arc/GIS Network Toolbox allows us to calculate the exact coordinates corresponding to the maximum GTC distance for the type of road and its specific attributes (capacity, gradient, congestion, etc.). Each GTC breakpoint represents an *isocost line* that lengthens as the capacity of the road network increases. We refer to these market isocost lines as *natural* trade areas, because they are defined in terms of *monetary* transport costs—not with proxies like distance—while relying on the statistical methods underlying Chow's structural test (i.e., confidence intervals).

These maps provide a first representation of the existence of an urban hierarchical system in terms of trade areas. From these maps several aspects emerge: *i*) the first breakpoint refers to the

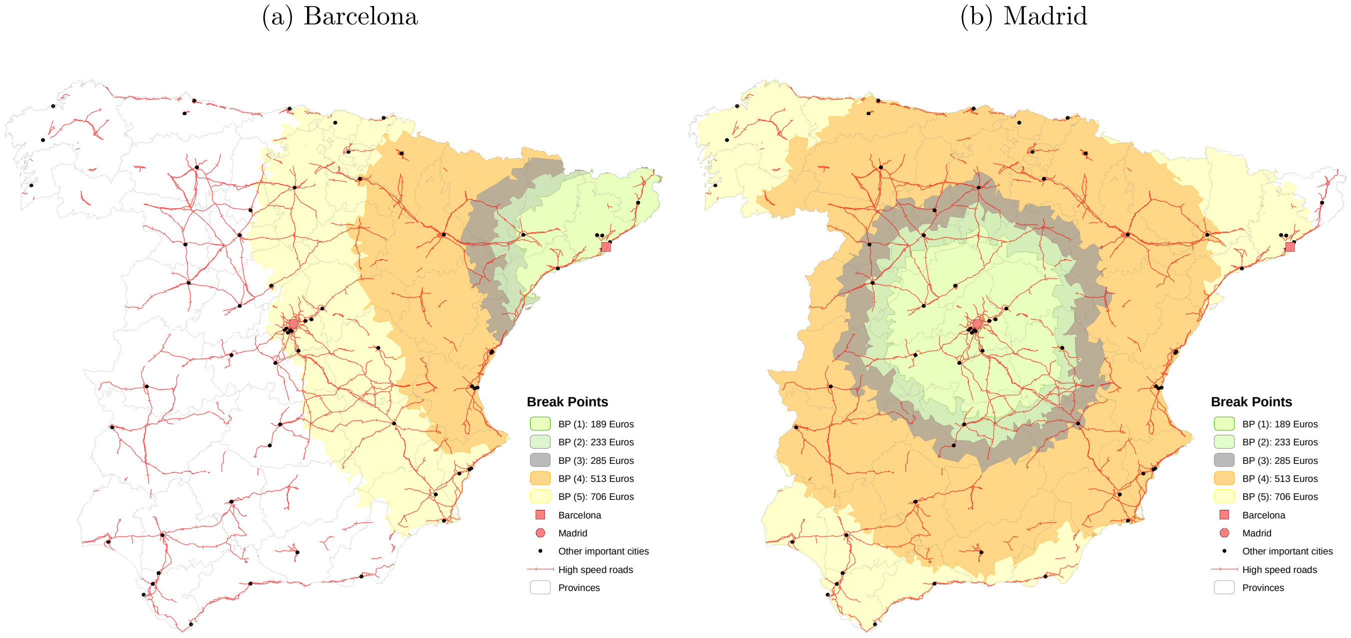
---

(2014)—see also Berthelemy and Varoudakis (1996) for an application of Chow's test in the economic growth literature.

[10] The GTC values in the maps are the same as those in Table 5 but expressed in distance-equivalent terms, so as to plot them correctly in the Arc/GIS software.

supply center between the main city and its metropolitan area and other, nearby satellite cities; *ii*) the second and third breakpoints successively reach some important cities (provincial capitals); *iii*) the fourth breakpoint appears as very relevant, as it joins Madrid and Barcelona with other Spanish cities that are wealthy in terms of trade (mainly Valencia and Zaragoza); and *iv*) the last breakpoint directly links Madrid and Barcelona, indicating that trade flows overlap for the two largest Spanish cities. Also, we observe that Barcelona and Madrid have stronger links to other cities densely surrounded by highways. This is the case especially with Madrid, whose road network-centrality allows for larger trade areas reaching all regions in Spain. Finally, note the geographical extension of the first breakpoint as it spans over municipal (NUTS-5) and provincial (NUTS-3) borders. As we can see, using administrative boundaries as a way to collect and infer internal border effects can be misleading.

Figure 2: Natural trade areas using GTC Breakpoints for the whole period data 2003-2007.


(a) Barcelona

(b) Madrid

With these findings we argue that once a correct definition of the dependent and explanatory variables—not proxies—is made available for the analysis, the methodology based on the Chow structural test is a promising procedure to determine trade areas. To our knowledge, there is not previous empirical valid measurement of cities' market areas based on trade flows and monetary transport costs (Löffler, 1998). More interestingly, as we argue in Section 5, these empirical regularities are in line with the predictions of the Lösch and Christaller's model, whose analytical framework constitutes the core of central place theory and gives rise to so-called urban hierarchical systems.

4.2. Accounting for GTC thresholds in the gravity model.

Once the breakpoints capturing structural changes in the trade-to-transport cost relationship are obtained, we split our GTC distance-variable according to these thresholds. We perform a set of regressions estimating Eq.(3) but eliminating the administrative borders for the interval-threshold in which, on average, no observation is recorded. For example, a transport cost interval of €285-€513 (breakpoints *3* and *4*) has no trade flows either at the intra-municipal level or between contiguous municipalities; these variables must therefore be dropped in a regression for this interval. We follow the same rationale for the rest of the transport costs intervals. In fact, we contend that this is the correct approach to capture the effects of transport costs and administrative borders in such highly granular trade flows as the ones we have computed. To lend support to this idea we include in the same table a general specification again using Eq.(3) and controlling for the quadratic non-linear GTC instead of the distance-thresholds approach. As we show in what follows, administrative borders overestimate the border effect over short distances; they have a different effect when we split the distance into thresholds.

Tables 6 and 7 show the PPML estimations by GTCs thresholds for the total value of trade and the extensive margin, respectively. The GTC is highly penalizing over short distances, but its negative impact on both types of flows decreases as distance increases. This confirms that the negative linear effect of trade frictions on trade flows varies over the spectrum of distances, but it also indicates that transport costs are not as detrimental to trade as normally thought. The impact of internal borders is very negative in both flows before the first threshold, especially the (municipal) NUTS-5 border as in Table 3. NUTS-3 and NUTS-2 again reduce their negative impact as GTC increases, except for the €513-€706 GTC interval (€582-€655 for the extensive margin), where the regional border has its highest impact on trade flows as a result of being the only border within the interval (extensive margin). Lastly, when GTC thresholds are not accounted for as in the "general specification," coefficients for trade frictions are distorted or even overestimated. We therefore conclude that standard gravity setups with granular trade flows poorly capture the (non linear) effects of transport costs and internal borders.

Table 6: PPML estimations by GTC thresholds. Total value of trade. Averages for 2003-2007.

	GTC Thresholds					General
Variables	(€0-€189)	(€189-€285)	(€285-€513)	(€513-€706)	(More €706)	Specification
GTC	-0.0145***	0.00262**	-0.00243***	-0.00329***	-0.00142***	-0.00525***
	(0.000339)	(0.00127)	(0.000185)	(0.000309)	(0.000125)	(0.000339)
GTC Square	—	—	—	—	—	3.69e-06***
						(2.47e-07)
Contiguity	0.911***	—	—	—	—	1.232***
	(0.0335)					(0.0702)
NUTS-5	3.022***	—	—	—	—	3.269***
	(0.0421)					(0.0836)
NUTS-3	1.059***	2.233***	—	—	—	1.247***
	(0.0348)	(0.318)				(0.0645)
NUTS-2	0.427***	0.0184	0.189***	1.347***	—	0.0945
	(0.0374)	(0.117)	(0.0708)	(0.102)		(0.0692)
FE Origin	Yes	Yes	Yes	Yes	Yes	Yes
FE Destiny	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,532	4,493	7,914	4,720	4,386	39,045
R^2	0.957	0.739	0.675	0.825	0.638	0.939

Clustered standard errors by *ij*. Std errors in parentheses. Significance level: ***p<0.01,**p<0.05, *p<0.1.
Note: The use of origin and destiny fixed effects and the lower number of observations in the (€189-€233) GTC interval forces to include Breakpoint 2 (€233) within the interval €189-€285.

Table 7: PPML estimations by GTC thresholds. Extensive margin. Averages for 2003-2007.

	GTC Thresholds						General
Variables	(€0-€185)	(€185-€246)	(€246-€321)	(€321-€582)	(€582-€655)	(More €655)	Specification
GTC	-0.0140***	-0.00357***	-0.00528***	-0.00175***	-0.00231***	-0.00133***	-0.00725***
	(0.000199)	(0.000572)	(0.000558)	(0.000143)	(0.000858)	(7.74e-05)	(0.000233)
GTC Square	—	—	—	—	—	—	4.48e-06***
							(1.64e-07)
Contiguity	0.767***	—	—	—	—	—	1.037***
	(0.0171)						(0.0359)
NUTS-5	2.556***	—	—	—	—	—	2.870***
	(0.0232)						(0.0518)
NUTS-3	0.808***	1.276***	1.553***	—	—	—	1.060***
	(0.0198)	(0.0789)	(0.101)				(0.0436)
NUTS-2	0.293***	-0.0164	0.161***	0.553***	—	—	0.118***
	(0.0231)	(0.0363)	(0.0487)	(0.0578)			(0.0423)
Origin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destiny FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,316	3,172	2,860	8,607	1,834	5,256	39,045
R^2	0.959	0.880	0.809	0.613	0.896	0.736	0.946

Clustered standard errors by *ij*. Std errors in parentheses
Significance level: ***p<0.01,**p<0.05, *p<0.1.

5. Trade areas and the hierarchical system of cities.

The GTC breakpoints and the accumulation of trade flows over short distances give rise to a series of empirical regularities matching the theoretical insights of central place theory (CPT) and its associated urban hierarchies à la Christaller. This theoretical framework is experiencing a revival thanks to a new series of studies (Parr (2002); Tabuchi and Thisse (2011); Hsu (2012); Mulligan et al. (2012); Hsu et al. (2014)). Based on CPT one expects areas of influence whose geographical reach is driven by transport costs; that is, consumers and firms 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 serve those locations for which consumers are willing to cover the transports costs of having the goods shipped to them. 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 where their market areas overlap. This predicted competition across locations and cities is shown in our data (Figures 2a and 2b).

In a stylized version of the CPT model, these cities' market areas are spatially represented by nested hexagons whose radii of influence are increasing in city's size and decreasing in transport cost. Whereas transport cost delimits the distance-breakpoint at which market areas change structurally, until they disappear, a city's size determine geographic extent of its commercial influence. Large cities can trade a large range of goods with many commercial partners, resulting in spatially larger GTC thresholds, whereas small cities trade few goods with a small number of cities. That is, in a hierarchy of cities à la Christaller we should expect few cities (Rank 1 cities) to present the largest market areas and supply a full range of products, a second group of large cities (Rank 2 cities) to serve a huge variety of commodities within a still relatively large market area, and a lower tier of medium-size and small cities (Rank 3 and Rank 4 cities) to be geographically scattered between cities of the two previous ranks, trading fewer products with lower or even insignificant market areas (Tabuchi and Thisse, 2011).

Empirical studies testing urban hierarchy systems have hitherto focused only on urban population patterns, even when theoretical predictions and the underlying assumptions relate it to the magnitude and composition of trade flows. Thanks to our highly granular trade flows, we can now test these predictions. Table 8 shows the trade distributions for the Spanish cities in 2003 and 2007, split by number of commodities traded and numbers of cities with which trade takes place. Data are given as percentage over the total number of municipalities (trading partners). As observed, the largest number of municipalities trade between 10 to 50 commodities with a set of 10 to 50 municipalities, even with a shift towards more commodities traded and trading partners in 2007. The same follows for upper intervals greater than 50 commodities and 50 municipalities. More interestingly these results shed light on a hierarchy of cities. The main diagonal characterizes various cities a vast majority of which are of small or medium size (Rank 4 and 3 cities), another group of cities trades an increasing number commodities with more cities (Rank 2 cities), and finally a group of only a few cities that trade a huge number of varieties with a large number partners (Rank 1 cities).

We get further evidence of the hierarchy of cities relying on cluster analysis, performed to consistently group cities by Christaller's predictions.[11] The classification of cities based on the analysis

---

[11] We base our cluster analysis on the two variables: the number of commodities and the number of trading partners. We use the K-means (centroids) algorithm and obtain four groups after 25 iterations.

Table 8: Shipments distribution by municipalities and products.

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 RFTS data.

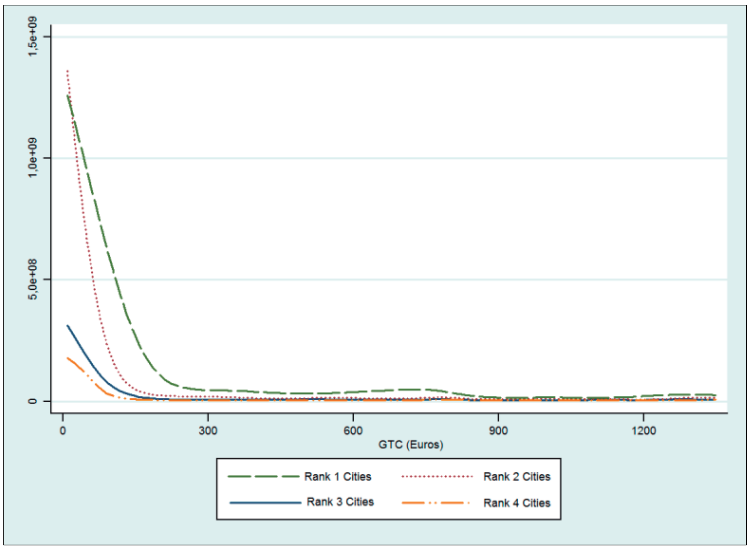
is reported in Table 9, with an overwhelming predominance of medium and small (Rank 3 and 4) cities and a skewness distribution for large (Rank 1 and 2) cities. Indeed, these city-groups should present different market areas depending on their rank. Figure 3 confirms these theoretical insights by way of kernel regressions for each cluster. Cities with a higher rank present higher densities for all distances (GTCs), especially over short ones and in the case of Rank 1 cities. Thus, the overall concentration of trade flows in short distances observed in Figures 2a to 2c is clearly driven by these Rank 1 and 2 cities. Notice also how these higher ranking kernels envelop those of lower city ranks.

Table 9: Number of Cities by Cluster.

City Rank	Number of Cities
Rank 1	4
Rank 2	21
Rank 3	101
Rank 4	507

We map the Spanish urban hierarchy from these results. To our knowledge, Figure 4 is the first illustration of a hierarchy actually based on trade flows, presenting cities in the four ranks. We also map a set of new breakpoints and market areas after applying the Chow test once again, but only for Rank 1 cities (Madrid, Barcelona, Valencia and Seville). The first threshold covers the metropolitan

Figure 3: Kernel regressions for different type of cities.



areas of these cities, while second and third breakpoints reach Rank 2 or even intermediate (Rank 3) cities, as predicted by CPT. Note that the geographical extent of these breakpoints exceeds even the internal administrative borders, particularly NUTS-5 (municipal) and NUTS-3 (provincial) territorial units; i.e., trading activity spans several administrative levels.

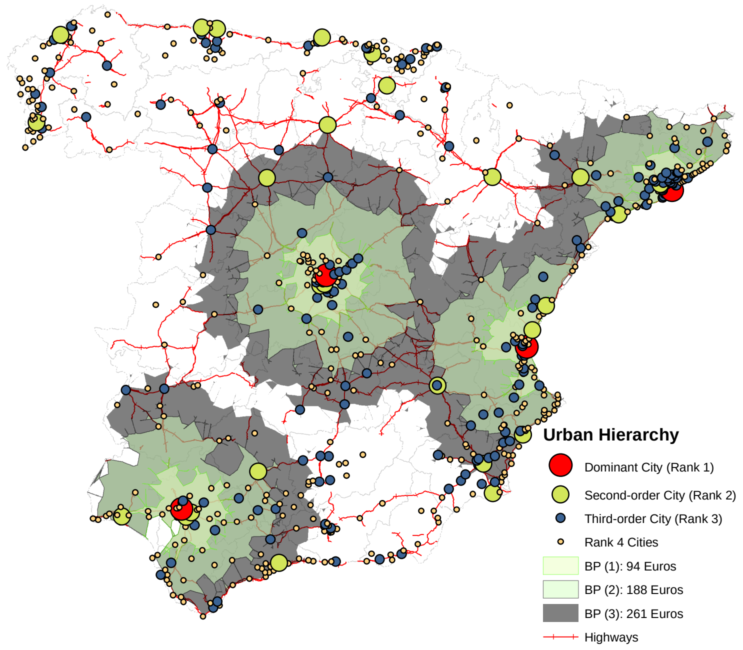
In light of these results we contend that the high agglomeration of trade flows at the municipal level in Table 3 is explained by the trading activity of Rank 1 and 2 cities, which in turn results in high market areas for these cities.

6. Conclusions.

We have compiled two unique and granular databases of municipal trade flows and monetary transport costs to assess trade agglomerations around specific areas; i.e., cities. We have also accounted for trade frictions and intra-national borders to show that regional borders have a much lower, or even negligible, impact on trade, whereas NUTS-5 (municipal) borders or surrounding areas concentrate the largest share of trade flows. In contrast to previous studies, we argue that this trade agglomeration over short distances has nothing to do with (unreal) border impediments, but arises as a result of transport cost thresholds that shape trade market areas around large locations, and thereby define a hierarchical system of cities.

To confirm this idea we have transposed the endogenous Chow test to the trade literature, allowing us to determine these specific trade-to-transport costs or breakpoints at which trade flows change structurally. With this approach, we have correctly measured the impact of internal borders and confirmed the high density of trade at GTC values below €189 (equivalent to 170km or 1 hour and 50 minutes). Finally, we have mapped these GTC breakpoints, whose analysis shows the emergence of an organized system of cities in terms of market areas. We argue that these results provide strong evidence for the insights of central place theory, in which trade agglomerations and reach are driven by the largest (dominant) cities within an urban hierarchy. All of these results call for future research related to the literature on urban patterns based on trade flows and transports costs.

Figure 4: Urban Hierarchy and Natural Trade Areas for Dominant Cities. First, Second and Third GTCs Breakpoints.



Note: The fourth breakpoint is equal to €655.

References

Abbate, A., De Benedictis, L., Fagiolo, G. and Tajoli, L., (2012), The international trade network in space and time. *LEM Working Paper Series*, N. 2012/17.

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

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

Anderson, J. and Yotov, Y., (2012). Gold standard gravity. *National Bureau of Economic Research*, Working Paper 17835.

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

Balistreri, E. and Hillberry, R., (2007). Structural estimation and the border puzzle. *Journal of International Economics*, 72 (2), 451-463.

Behrens, K., Mion, G., Murata, Y. and Suedekum, J., (2017). Spatial frictions. *Journal of Urban Economics*, 97, 40–70.

Benedictis, L. and 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, 55–89.

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

Borraz, F., Cavallo, A., Rigobon, R. and Zipitria, L., (2016). Distance and political boundaries: Estimating border effects under inequality constraints. *International Journal of Finance and Economics*, 21(1), 3–35.

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

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

Crafts, N. and Klein, A. (2014). Geography and intra-national home bias: US domestic trade in 1949 and 2007. *Journal of Economic Geography*, 15(3), 477–497.

Diallo, I., (2014). Suchowtest: Stata module to calculate successive chow tests on cross section data. Statistical Software Components S457536, Boston College Department of Economics, revised 23 Jan 2014.

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

Gallego, N. and Llano, C., (2014), The Border Effect and the Nonlinear Relationship between Trade and Distance. *Review of International Economics*, vol. 22 (5), pp. 1016-1048.

Gallego, N., Llano, C., De la Mata, T. and Diaz-Lanchas, J., (2015). Intranational Home Bias in the Presence of Wholesalers, Hub-spoke Structures and Multimodal Transport Deliveries. *Spatial Economic Analysis*, 10 (3), 369- 399.

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

Head, K., Mayer, T., (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, Elsevier, Vol. 4, pp. 131-195.

Hillberry, R., (2002)a. Aggregation bias, compositional change, and the border effect. *Canadian Journal of Economics*, 35 (3), 517–530.

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

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

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

Hsu, W., (2012). Central place theory and city size distribution. *The Economic Journal*, 122 (563), 903–932.

Hsu, W., Holmes, T. and Morgan, F., (2014). Optimal city hierarchy: A dynamic programming approach to central place theory. *Journal of Economic Theory*, 154, 245-273.

Llano, C., Esteban, A., Pérez, J. and Pulido, A., (2010). Opening the interregional trade "black box": The c-intereg database for the spanish economy (1995-2005). *International Regional Science Review*, 33 (3), 302–337.

Llano, C., De la Mata, T., Díaz-Lanchas, J., and Gallego, N. (2017). Transport-mode competition in intra-national trade: An empirical investigation for the Spanish case. *Transportation Research Part A: Policy and Practice*, 95, 334-355.

Llano-Verduras, C., Minondo, A. and 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.

Martínez-San Román, V., Mateo-Mantecón, I. and Sainz-González, R., (2017), Intra-national home bias: New evidence from the United States commodity flow survey, *Economics Letters*, 151, pp. 4-9.

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

Mori, T. and Wrona, J. (2018). Inter-city Trade. *CESifo Working Papers*, 7233, ISSN 2364–1428.

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

Nitsch, V. and Wolf, N., (2013). Tear down this wall: On the persistence of borders in trade. *Canadian Journal of Economics*, 46 (1), 154-179.

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

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

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

Ramos, R. and Moral-Benito, E., (2017). Agglomeration by export destination: evidence from Spain. *Journal of Economic Geography*, 18 (3), 599-625.

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

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

Santos Silva, J. and 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., (2011). A new economic geography model of central places. *Journal of Urban Economics*, 69 (2), 240–252.

Thissen, M., Di Comite, F., Kancs, d'A. and Potters, L., (2014). Modelling inter-regional trade flows data and methodological issues in RHOMOLO. *European Commission Working Papers Series*, 02/2014.

Wolf, H., (1997). Patterns of intra-and inter-state trade. *National Bureau of Economic Research*, Working Paper 5939.

Wolf, H., (2000). Intranational home bias in trade. *Review of Economics and Statistics*, 82 (4), 555-563.

Wrona, J., (2018). Border effects without borders: What divides Japan's internal trade? *International Economic Review*, 59 (3), 1209-1262.

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

Yilmazkuday, H., (2011). Agglomeration and trade: State-level evidence from U.S. industries. *Journal of Regional Science*, 51 (1), 139–166.

Yilmazkuday, H., (2012). Understanding interstate trade pattern. *Journal of International Economics*, 86, 158–166.

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

Supplemental online appendix – Not for publication.

A.1. A second level decomposition of the extensive and intensive margins.

The extensive and intensive margins in Eq.(1) can be further decomposed, so the total number of shipments (N_{ij})—extensive margin—can be divided into the number of distinct commodities traded (k) between each pair ij (N_{ij}^k), and its frequency or average number of shipments per commodity (N_{ij}^F):

$$N_{ij} = N_{ij}^k N_{ij}^F. \tag{A.1}$$

In turn, the intensive margin is split into average price ($\overline{P}_{ij}$) and average quantity ($\overline{Q}_{ij}$), where s indexes unique shipments:

$$\overline{PQ}_{ij} = \frac{(\sum_{s=1}^{N_{ij}} P_{ij}^s Q_{ij}^s)}{N_{ij}} = \frac{(\sum_{s=1}^{N_{ij}} P_{ij}^s Q_{ij}^s)}{(\sum_{s=1}^{N_{ij}} Q_{ij}^s)} \frac{(\sum_{s=1}^{N_{ij}} Q_{ij}^s)}{N_{ij}} = \overline{P}_{ij}\overline{Q}_{ij}, \tag{A.2}$$

Additionally, we also consider regression results based only in physical units (tons), leaving prices aside:

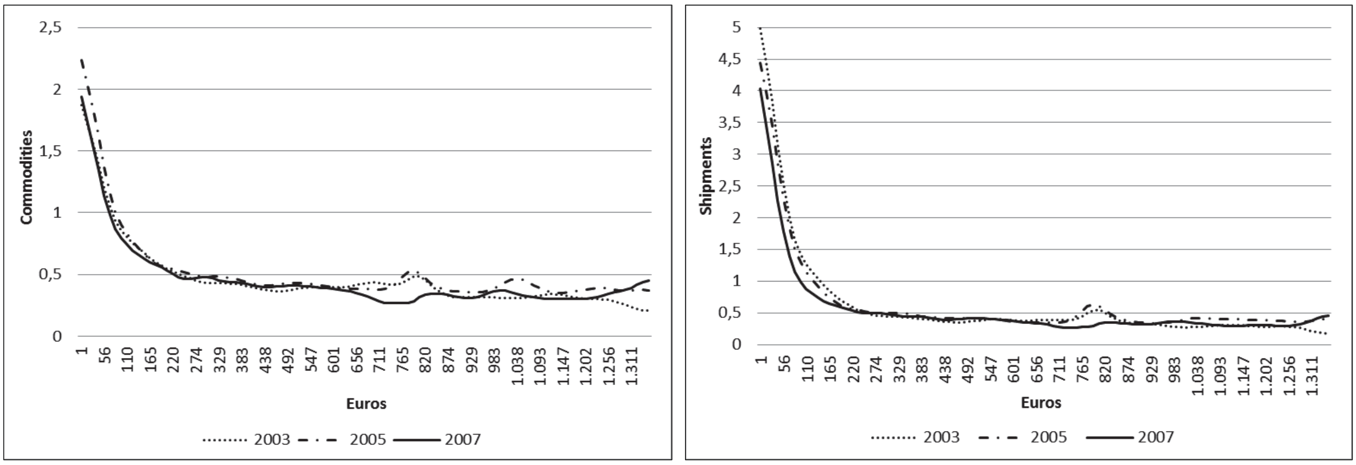
$$Q_{ij} = (\sum_{s=1}^{N_{ij}} Q_{ij}^s) = N_{ij}\overline{Q}_{ij}. \tag{A.3}$$

In the database we follow a bottom-up approach such that, to obtain Eq.(A.1), we calculate the maximum number of different commodities between each ij pair and multiply it by its frequency. Apart, we calculate average quantity ($\overline{Q}_{ij}$) and average price ($\overline{P}_{ij}$) across all shipments between ij, thereby obtaining an average value per shipment ($\overline{PQ}_{ij}$); i.e., Eq.(A.2). Then, we multiply it by total number of shipments (N_{ij}) to get total value of trade as in Eq.(1). For the total trade in quantities Eq.(A.3), we multiply the extensive margin, Eq.(A.1), by the average quantity.

A.2. The density of the trade components of the extensive and intensive margins.

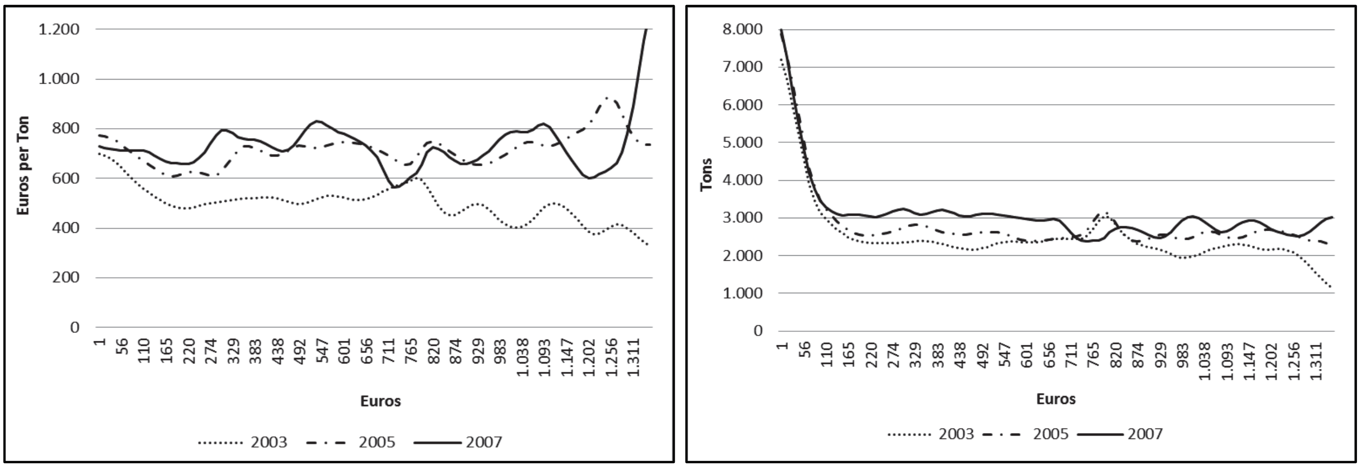
Figure A.1 shows the kernel regressions for the second decomposition. Number of commodities (Fig.A.1a) and its frequency (Fig.A.1b) exhibit a remarkable similar spatial pattern as the extensive margin, i.e., they sharply fall with the GTC. Average price (Fig.A.1c) presents an increasing trend in transport costs, whereas average tons (Fig.A.1d) are again highly concentrated in short distances and become flat in medium and longer ones.

Figure A.1: Kernel regressions of the extensive margin (second decomposition).



(a) Number of commodities

(b) Number of Shipments per Commodities



(c) Average Price

(d) Average Tons

Figure A.2: Kernel regressions of the intensive margin (second decomposition).

A.3. PPML regressions for the extensive and intensive margins decomposition.

Table A.1 replicates the regressions performed in Table 3 for the extensive and intensive margins. Number of shipments per commodity (frequency) and number of different commodities shipped explain approximately the same proportion of their margin for all administrative boundaries, especially for the municipal border (NUTS-5). NUTS-3 and NUTS-2 variables exhibit the same pattern that in the first level decomposition; i.e., provincial (NUTS-3) borders reduce their importance as trade deterrents while regions (NUTS-2) have an ever lower impact on trade flows. Average tons shipped within the municipality are higher than inter-municipalities ones, although they show a decreasing trend between borders, and even exhibit a negative impact at the regional level (NUTS-2). Finally, we confirm the robustness of our coefficients when considering physical quantity measures, even more for regional borders (NUTS-2), which are not statistically significant.

Table A.1: PPML estimations with GTC and internal borders (second decomposition).

	Extensive margin		Intensive margin		
Variables	N. of Commodities (N_{ij}^{K})	Trading Pairs (N_{ij}^{F})	Price ($\overline{P}_{ij}$)	Tons ($\overline{Q}_{ij}$)	Trade in Quantities (Q_{ij})
GTC	-0.154***	-0.182***	-0.00772	-0.114***	-0.447***
	(0.00467)	(0.00566)	(0.00469)	(0.00420)	(0.0292)
GTC Square	4.17e-11***	6.47e-11***	-1.09e-11***	4.41e-11***	1.88e-10***
	(2.50e-12)	(2.89e-12)	(2.77e-12)	(2.28e-12)	(1.50e-11)
Contiguity	0.659***	0.827***	0.123***	0.665***	1.262***
	(0.0141)	(0.0264)	(0.0232)	(0.0222)	(0.0518)
NUTS-5	1.507***	1.933***	0.504***	1.122***	3.495***
	(0.0333)	(0.0371)	(0.0435)	(0.0388)	(0.0792)
NUTS-3	0.845***	0.781***	0.432***	0.438***	1.415***
	(0.0169)	(0.0206)	(0.0208)	(0.0178)	(0.0655)
NUTS-2	0.239***	0.139***	0.188***	-0.0321*	0.0767
	(0.0168)	(0.0207)	(0.0207)	(0.0166)	(0.0554)
Year FE	Yes	Yes	Yes	Yes	Yes
Origin FE	Yes	Yes	Yes	Yes	Yes
Destiny FE	Yes	Yes	Yes	Yes	Yes
Observations	195,026	195,026	195,026	195,026	195,026
R^2	0.076	0.146	0.512	0.344	0.624

Clustered standard errors by *ij*. Std errors in parentheses. ***p<0.01,**p<0.05, *p<0.1.

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub
EU Science Hub - Joint Research Centre
Joint Research Centre
EU Science Hub