Trade and spatial agglomeration: how spatially concentrated are inflows in the importing country?

Nuria Gallego*
Carlos Llano**

*Departamento de Economía Aplicada I, Historia e Instituciones Económicas y Filosofía Moral. Facultad de Ciencias Jurídicas y Sociales, Universidad Rey Juan Carlos, Campus de Vicálvaro (Madrid).

**Departamento de Análisis Económico: Teoría Económica e Historia Económica. Facultad de Ciencias Económicas y Empresariales. Universidad Autónoma de Madrid. Campus Cantoblanco. 28049 MADRID.

Abstract

In a world of heterogeneous space (asymmetric regions), some areas like border regions benefit from geographical advantage when trading with foreign markets. To date, the literature has concentrated mainly on analysing the internal effects of openness to trade, by measuring to what extent the liberalisation process triggers the internal agglomeration of firms and jobs close to the border regions. By contrast, little attention has been devoted to the destination side of the issue: the agglomeration of the flows in the importing country. In this paper we explore the presence of different spatial pattern in the concentration of the inflows in the importing regions when the exports come from the border regions or from any other region in the exporting country. By doing so, we wonder if the spatial pattern of border regions are mainly explained by this geographical advantages (accessibility). The empirical analysis uses a novel dataset on interregional flows between the Spanish regions and the regions (Nuts 2) of other 8 main European countries. To the best of our knowledge is the first time this analysis has been developed, mainly due to the lack of information regarding inter-regional inter-national bilateral flows in Europe.

JEL Codes: F14, F15
Keywords: interregional trade, region to region international trade, accessibility, border regions, distance.
1. Introduction

Some recent papers have investigated the relationship between international trade performance, firm location choices and the geographical advantages associated with coastlines, trade gateways or certain transport infrastructures (Bleakley & Lin 2012, Duranton et al., 2012; Atkin and Donaldson, 2012; Cosar and Fajgelbaum, 2013; Fujita and Mori, 1996). In a related literature, different authors have considered how accessibility to international borders affects the international concentration of the economics activity within a country (Brülhart 2011, Redding and Sturm, 2008; Baldwin and Wyplosz, 2006; Nitsch, 2006; Rodriguez-Pose, and Gill, 2006; Brülhart and Traeger, 2005; Mansori, 2003; Hanson, 2001), usually considering the fact of a sudden trade liberalisation as a quasi-experiment.

According to Brülhart (2011), neither the theoretical models, nor the empirical research are conclusive on whether an increase in the openness of a country has a clear impact on the internal reallocation of firms and workers, and therefore, on an increase in the internal disparities in terms of wages or per capita income. In despite of this indeterminacy, it seems more plausible that, when the heterogeneity of space is considered (asymmetric regions), trade liberalisation increases the spatial inequality within a country, in favour of the regions with better access to the new markets. To this regard, it is argued that “port locations and navigable rivers are evident examples of natural features that facilitate access to distant markets. In modern economies, however, market access is shaped to an ever larger extent by man-made infrastructure, including roads, railway links, airports and telecommunication networks”. Then, the survey concludes suggesting a list of potential fruitful directions for future research, to say: the analysis of the differential impacts of reductions in trade costs that are related or independent of distance (tariff versus transport costs); the analysis of the explicit causal relation between the impact of trade on intra-national economic geographies (by exploiting quasi-experimental settings); and the consideration of industry specific mechanisms enhancing the spatial effects of trade liberalisation.

Most part of the empirical papers revisited tend to explain how liberalization promotes the agglomeration of economic activity (firms, production and jobs) in the regions facing geographical advantages with regard to the international markets. Thus, from the empirical perspective, the usual approach tries to test whether the spatial concentration of firms within the exporting country (at the regional or urban level) has increased along with the level of openness and integration with the external markets. As a recent example for the Spanish case, Ramos. & Moral-Benito (2013) used a unique dataset of Spanish exporters to document the existence of exporters’ geographical agglomeration by export destination, finding that firms selling to countries with worse business regulations, a dissimilar language and a different currency tend to cluster significantly more. Such results were obtained considering agglomeration in a model of international trade à la Melitz (2003), using Spanish firm-level data.

In contrast, this literature has paid little attention to the spatial concentration of the flows themselves, that is, considering the spatial clustering of one’s export flows into the importing regions in the destination country. This is exactly our main contribution to this literature. The goal of our paper is to search for different
agglomeration patterns of the imports in the destination country, one control for the exporting region and the nature of the products being delivered. Our intuition is that border regions have a larger propensity to export to the adjacent regions in the foreign country, or in other words, that the spatial pattern of flows delivered from the border regions in the exporting country is mainly driven by the geographical advantages (accessibility to the nearest regions in the importing country) rather than the market size (the wealthiest regions in the importing country).

In this paper we explore a real case of asymmetric regions and heterogeneous space in Europe. Our analysis centres on the actual interregional trade flows delivered by road between Spain and 8 main European partners, with a special consideration to the Spanish-French border, the main gateway to the rest of the continent. As we will describe later on, this specific border is of great interest for our purpose, considering its historical, economic and geographic specificities.

Thanks to a singular dataset, which covers interregional flows from the Spanish regions to the regions in the 8 main European partner countries (Gallego and Llano, 2012, 2014, 2015), we are able to investigate the spatial patterns of the exports taking into account the actual trading regions at both sides of the national borders. First, by means of kernel regressions, spatial autocorrelation indexes and cluster analysis we are able to identify different patterns of spatial concentrated/disperse trade, depending on the geographical advantage of some regions, which are located close to the main gateways from Spain to the continental Europe through the French border (Girona and Irun). Then, by means of the gravity equation, we are able to test if the spatial concentration of the trade flows is statistically different depending on the region’s actual travel time to the French border.

To this regard, the results obtained are complementary to the ones published by other authors, revelling that the countries located contiguous to the national borders show higher levels of concentration of trade with the nearest regions in the neighbouring countries. Our investigation also reveals an interesting geographical specialisation of the Spanish regions when exporting to the European regions, which is also driven by their relative geographical advantage with regards to the eastern or western coast of France. To the best of our knowledge, this paper is the first one analysing the concentration of trade flows with a region-to-region dimension and putting it in the context to the accessibility to the international borders in Europe.

The structure of the rest of the paper is the following: section 2 review the literature that have analysed the relation between trade openness and the internal allocation of the economic activity. Then, section 3, describes the methodology used for estimating a region-to-region trade dataset for the Spanish case. In section 4, we develop different quantitative analysis with the aim of identifying spatial concentration patterns of the flows themselves, both considering aggregate Spanish exports and imports, as well as industry specific flows with the main European partners.
2. The empirical strategy.

The aim of our empirical exercise is analyse the presence of different spatial patterns in the intensities of imports arriving to the regions of an importing country (i.e. France), once that one control for the geographic characteristics of the exporting regions in the country of origin (Spain), as well as the nature of the product being delivered. To this regard, we wonder whether border-regions have different spatial patterns than non-border ones, and to what extent exports from border-regions of the exporting country are spatially clustered in the nearest regions in the importing country, and not necessarily in the richest ones. This analysis is then expose to additional robustness checks, considering that the key factor is not just being or not a border region in the exporting country, but being at certain distance to the main gateways to the foreign market.

In order to illustrate our intuition, Figure 2 describes two archetypical patterns of trade, considering, for example, the Spanish case. On the one hand, according to panel A), labelled “trade costs driven trade”, it is expected to find higher spatial concentration of trade in the regions located close to the border in the exporting country, which have generated/attracted -through History- the firms with a larger propensity to trade in shorter distance to the neighbour country (i.e.: red shadowed areas corresponding to border regions such as País Vasco or Cataluña). We then want to test if a statistically significant larger level of imports also concentrates in the closest regions to the border at the other side of the border. By contrast, Panel B) describes an alternative pattern of trade, labelled as “market size driven trade”, where the largest exports with origin in a region (i.e. Madrid) do not concentrate in the closest regions in France, but in other regions (i.e. Paris, Brussels, Munich, Rome…), whose market size, or some idiosyncratic features, compensate for their lack of accessibility compared to the regions located close to the borders.\footnote{Although the example suggested here is connected with the reality described thereafter in our data, it is important to avoid any positive/negative connotation regarding the two patterns of trade described here. Neither of both is better than the other. In fact, one may argue that in the case of A) border regions are enjoying location advantages derived from geography (first and second nature), while in the case of B) Madrid is benefiting from all the privileges derived from being the capital of the country, the main location for the headquarters of the main firms in the countries, and the “node” of a centralized-transport-network. Moreover, as we will see in our empirical investigation, having a larger agglomeration of trade in the nearest regions do not impede to also have intense flows with further markets. Furthermore, note that our investigation just considers deliveries by trucks. Therefore, all the deliveries to further markets using other transport modes are excluded from this analysis.}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Archetypical patterns of trade, considering the Spanish case.}
\end{figure}

The two labels described before are defined in order to remark two potentially different patterns of internationalisation, which in fact would be conditioned by a number of industry specific factors such as the transportability of the products, the favourite transport mode, the level of competitors and the presence of social and business networks. Obviously, in reality none of these pure cases usually hold.

2.1. A formal model.

The gravity model is the workhorse of many empirical analyses in international trade. According to its original definition (Tinbergen, 1968), the bilateral trade between any two locations emulates the attraction force of two
masses in the space. This notion of gravity is rather particular, since it focus on the bilateral attraction force between 2 objects located in the outer space, with no other constrain than the distance. Our approach could be described as an enriched gravity model, where the pull factors \((Y_i)\) and the trade costs \((d_{ij})\) plays a major role for some regions (those with geographical advantages, that is, with the best accessibility to foreign market). For these regions, trade will spillover to closest regions, *swamping* these markets with a larger intensity than what a normal gravity model will predict. Thus, for these border regions an abrupt decay of trade with the distance after crossing the border \((dist_{tf})\) is expected. Moreover, a positive spatial autocorrelation on the inflows arriving to the most accessible regions in the importing country is also probable.

Formally, we express such intuition in the following empirical form:

\[
\frac{e^{eu}_{ij}}{Y_iY_j} = f\left(dist_{tf}; dist_{fj}; dist_{tf}^2\right) \quad [1]
\]

Where \(\frac{e^{eu}_{ij}}{Y_iY_j}\) is the intensity of bilateral flows between two regions \(i\) and \(j\) in two different countries \((e=Spain\) for Spanish exports) relative to the GDP’s of the exporting and importing regions. \(dist_{tf}=\) distance from the exporting region to the most convenient entrance through the French-Spanish border. \(dist_{fj}=\) distance from the most convenient entrance through the French-Spanish border to the final destination in the importing country. These variables are explained in more detail in the next sections.

Similarly, we expect that for some regions (border-regions mainly) the exports in the importing country will be spatially agglomerated in the group of foreign regions located closer to the frontier:

\[
\frac{e^{eu}_{ij}}{Y_iY_j} = f\left(dist_{tf}; dist_{fj}; dist_{tf}^2; \rho W\left(T^{eu}_{ij}/Y_iY_j\right)\right) \quad [2]
\]

\(\rho W(T^{eu}_{ij}/Y_iY_j)\) = with this general element we denote the spatial association of intense inflows entering in the neighbouring regions of any foreign region \(j\). We expect to find the higher positive and significant \(\rho\) coefficients the lower \(dist_{tf}\) (border regions in the exporting country) and the lower the \(dist_{fj}\) (border regions in the importing country).

[This section, and the formal expression of the models, have to be improved]

3. The data

3.1. Some relevant geographical and historical remarks

The French-Spanish border is a line of great interest for anyone willing to consider the effect of geography (first nature) and history (second nature) on trade and the actual allocation of the economic activity. In our view, it offers the required elements to define a quasi-experiment regarding the issues described in this paper.

Historically, the Spanish-French border has been a field of controversy and intense military interactions\(^2\). Spain and France has been immersed in open confrontations, or enrolled in opposed alliances for most part of

---

\(^2\) An in depth review of the history of the Spanish-French border is out of the scope of this article. Useful information can be obtained in specialized references Maland (1991) Pendrill (2002) as well as different voices in Wikipedia.
their History. In that sense, until the end of the II World War, although there has been some positive connections between both countries (i.e. many centuries of common royal families), this border can be easily described as a great political obstacle for economic interactions. The Spanish-French border extends over 656.3 kilometres southwest of France and northeaster Spain. It begins in the west over the Bay of Biscay (Irun), and continues eastward along the Pyrenees to Andorra, and then continues east to the Mediterranean Sea (Girona). The creation of the border dates to the Pyrenees Treaty (1659). The current trace was set in Bayona Treaty (1868) with little modifications in 1980. The Treaty of Pyrenees was signed between France and Spain in the context of the Thirty Years’ War, which allowed France to gain Roussillon and Perpignan (both in the current French region of Languedoc-Roussillon). However, Spain still preserves the small town of Llívia in this part of France, which still belongs to Girona (Cataluña). On the western side of the border there are also strong historical connections between the Spanish and French border regions (Pais Vasco-Aquitania).

**Economically**, the Iberia peninsula could be seen as a ‘cul de sac’ for the rest of Europe, but also as a strategic enclave for the interrelation with Africa and America. On the contrary, from the Spanish side, the French-border is the gateway to the European core, the main Spanish market both for exports and imports. As a consequence of the different economic effect of crossing the border to reach one or the other market, one may expect to find asymmetric effects in the spatial pattern of the flows entering from Spain into France and vice versa. Moreover, one may also want to consider that due to the North-South dichotomy existing in many EU countries, it is here the case that some of the more industrialized regions in Spain are located in the north (Cataluña and País Vasco) while the regions in the south of France are among the less developed in that country (mainly on the south-west and south-centre part of France). Without the intention of disentangling the endogeneity of this situation and its relative position regarding the European core, it is probable that such dichotomy will also play a role on explaining asymmetric spatial patterns of concentration for the observed trade between the two countries.

**Geographically**, this border is also singular. The frontier runs along the Pyrenees, a range of abrupt mountains that extends for about 491 km, from the Mediterranean coast to the Cantabric Sea with more than 50 picks with an altitude above 3,000 meters and more than 30 between 2,999 and 1,000 meters. The whole border (656.3

---

3 It could be the case that a number of highly productive firms have agglomerated in the Spanish northern regions close to the French border, with the aim of enjoying location advantages for trading with the rest of Europe. Then, when they try to enter in the southern French regions, they find a lower level of competition than in other regions in the north or France, where it is more likely to find the most productive French regions willing to be as close as possible to the European core. This effect could be clearer for some industries than for others. In fact, it could be critical for products with low levels of transportability. Let us consider that the firms from certain industries in France have also tended to locate close to their strategic markets (EU Blue Banana), which in this case are in the North (close to Germany) and along the Mediterranean coast, rather than close to the south-western French border with Spain. Thus, it could be the case that the firms located in the French south-western regions will be weaker when competing with more competitive Spanish firms, which have been agglomerated close to the border. In fact, the Spanish firms located close to the French border could serve the French southern regions even with lower transport costs than any French competitor located in the north of France. Simply, it could be the case that some southern French locations are marginal markets for highly competitive firms located close to the Blue Banana, while these same regions are an interesting gateway for the Spanish strongest exporters. Furthermore, it is important to make notice that it is not possible to make the difference between Spanish firms and foreign firms located in either of both sides of the border. To this respect, it could be the case that French firms interested in the Iberian market have opt for investing in Spanish regions (i.e. large FDI outflows from France and Germany entering in some Spanish industries such as the automobile, the distribution or retail sector), serving both southern French regions, and the whole Spanish market from Spain itself. If this were the case, part of the cross-border trade flows observed in the data could just corresponds to trade flows generated by foreign affiliates located in Spain.
km) just offers 2 points of access by highway through Irun (Pais Vasco) and Girona (Cataluña). Although there are other roads connecting the two countries, most part of the transport flows occur thru these two corridors, which indeed perform as gateways for the main economic interactions by ground transportation. The use of alternative transport modes has been almost impossible during many years. A well-known example of this is the development of an exclusive ‘Iberian gauge’ for the Spanish and Portuguese railways (the distance between the two rails of a track is 1,668 mm), which is different to the most common European standard train gauge (1,435 mm). Such peculiar Iberian standard was set in 1845 and then ratified in 1955, with the aim of avoiding French invasions. In the 90’ new high-speed passenger lines in Spain were built using the international standard gauge, which assures the connectivity with the European high-speed network.

3.2. A region-to-region international trade dataset for Spain.

Currently there is no official data on region-to-region international trade flows for any country in the EU. Gallego and Llano (2012) describes a methodology for estimating region-to-region international flows between Spain and 8 European countries, by combining region-to-region freight statistics on Spanish trucks \( e_{ugij} \), as well as international price indexes for each region-country-variety \( e_{ugij} \) deduced from official trade data. The outcome of this procedure is a unique dataset with information on region-to-region flows for the international flows departing/arriving to the Spanish regions (Nuts 2) and the regions of the 8 main European partners. Each flow can be described by equation [3]:

\[
\hat{e}_{ugij} = \left( e_{uj}^R, \hat{e}_{ugij} \right)
\]

The estimated trade flows \( \hat{e}_{ugij} \) of products \( g \) traveling by Spanish trucks from region \( i \) in country \( e \) to region \( j \) in country \( u \), were obtained by the combination of the estimated trade prices \( \hat{e}_{ugij} \) (or \( \hat{e}_{ugij} \) for imports, being \( u=\text{Spain} \)) and actual freight flows in tons, \( e_{uj} \), delivered by Spanish trucks. Note that if \( e=\text{Spain} \), equation [3] will capture Spanish exports to regions in our sample of 8 European countries, while if \( u=\text{Spain} \), it will capture the Spanish imports from the regions of such countries (the list of countries and products is reported in the Annex). Due to the characteristic of our road flow dataset, we exclude flows where a Spanish island-region is a partner.

4 Although for simplicity we use the label EU, the sample of countries considered here does not specifically fulfill any administrative category, and indeed includes countries such as Andorra, a single-region country.

5 In most of the EU countries, there are two main sources on bilateral flows of goods between countries: 1) Trade statistics on intra-EU trade, which register bilateral flows between pairs of countries, both in volumes and monetary units; in some countries like Spain, trade data identify the exporting region or the importing region, but never both simultaneously; 2) Transport statistics on intra-national and inter-national freight flows, which -in some cases (road freight)- contain information on the type of product transported (just in volume), as well as on the specific regional origin and destination of the flows. The methodology implemented here aims to build up a region-to-region trade dataset by combining these two sources: 1) region-to-region flows in volume (road freight statistics); 2) region-to-country specific trade prices (from the official trade statistics).
Regarding the distance variable, we use the mean *actual distance* covered by the Spanish trucks in the operations between each pair of trading regions. This data was also reported by the same survey published by the Spanish Ministerio de Fomento. This measure has the virtuosity of capturing the *real distance* travelled by trucks between actual origins and destinations. Departing from this information for each delivery in the period (2004-07) we built a variable $\text{dist}_{ij}$ by averaging the distance traveled for each $i$-$j$ pairs during the period. Then, we want to split this total distance in two parts, considering the distance from the exporting region $i$ to the frontier $f$, and from there to the importing region $j$. Unfortunately our transport survey do not include any information regarding the route followed by the trucks when crossing the Spanish-French border, or the distance actually travelled in both sides of the border. However, thanks to the singular geography of that border, which have two main entrance at both extreme locations of the border (Girona in Cataluña and Irun in the País Vasco), we could easily define an *accessibility* measure to these main corridors. Thus, by means of the actual Spanish road network and a Geographical Information System (GIS), we compute an accurate optimal distance from each Spanish region to these two strategic enclaves. More specifically, for each of the Spanish regions (Nuts 2), we have computed an *optimal travel distance* to Girona and to Irun, which have been obtained as a weighted average of the corresponding *optimal distance* from the capital provinces (Nuts 3) in each region (weighted by their population) to Girona and Irun. With that, we assume that $\text{dist}_f$ for the Spanish exports (and $\text{dist}_f$ for the Spanish imports) is equal to the computed optimal distance travelled from the Spanish region ($i$) to the closest main gateway in the French border ($f$); while the $\text{dist}_f$ for the exports of each of the Spanish regions (and the $\text{dist}_f$ for the imports) is obtained as a residual between the actual-total and the computed optimal distance to the border ($\text{dist}_f = \text{dist}_{ij} - \text{dist}_f$).

In addition, by means of the same GIS and some detailed information obtained from an alternative research project (www.uam.es/Transporttrade/) we were able to obtain *“minimum travel times”* and *“minimum generalized transport costs”* (Zofío et al, 2011) by road also from each Spanish province (Nuts 3) to the two main gateways to Europe through the French border. The corresponding measures at regional level (Nuts 2) are obtaining by a weighted average of the optimal measures at the province level (Nuts 3). Note that equivalent information for the European network is not available. Therefore, *optimal travel times* and *transport cost* will be just computed for the internal movements within Spain. For this reason, such variables are just used as robustness checks in section 5.1.2 but not in 5.1.3. Finally, GDPs and other main indicators for each region are taken from Eurostat.

### 4. Empirical analysis

In this section we analyse the spatial patterns of the Spanish international trade with the regions of 8 European countries. Although the analysis focuses on the Spanish case, and is conditioned by the economic and geographical features of the country, it is developed with the aim of answering the general questions posed above, trying to test the following hypothesis:
i. Focusing on Spain as an exporter, the accessibility to ‘the European Core’ is a source of ‘geographical advantage’ with a heterogeneous impact on the competitiveness of the Spanish regions. Therefore, the northern Spanish regions located close to the national border with France (gateway to the main European markets, neglecting Portugal) may exhibit lower transport costs, both for exports and imports, and therefore, could play a role as ‘hubs’ for the rest of Spain-Europe (Ghemawat et al., 2010; Llano and Díaz, 2011; ). For this reason, a larger concentration of exports to Europe -as well as import flows from Europe- are expected;

ii. If the specific geography of a country matters, actual travel times and transport cost, and not just the fact of sharing a border, is the real driver for achieving geographical advantages when exporting abroad. To this regard, transport costs and travel times, and not straight line distances will determine the accessibility and geographical advantage. To this regard it is important to consider not just the accessibility from the national border to the importing regions, but also the transport cost within the country (from the exporting region to the frontier). Thus, although there are 4 Spanish regions located along the border with France, the Pyrenees clearly reduce the accessibility of two of them (Aragón and Navarra) compared to the other two located in the coast (País Vasco and Cataluña), which are crossed by highways, and canalized most of the international trade flows by road (Irun vs Girona);

iii. In spite of current “geographical advantage”, historically, industrial firms (Spanish exporting companies as well as multinationals) may have tended to locate close to the French border (in both sides) looking for better accessibility. Therefore, one may expect to find larger intensities of trade between these regions (mainly for the case of Cataluña-Languedoc-Roussillon, País Vasco-Aquitaine; and less clear for Aragón-Navarra-Midi-Pyrenees). This effect corresponds to the cumulative process of concentration of exporting firms near the border described in the literature (Brülhart, 2011).

Departing from these hypotheses, we develop an empirical analysis of the novel dataset described before combining a set of methodologies. The analysis is applied to the Spanish exports using the following tool-kit:

i. First, by means of kernel regressions we analyse the distribution of trade as regards to distance. By doing so we expect to identify homogeneous groups of regions, regarding the distribution of their trade flows with the rest of Europe;

ii. Second, by means of different indexes of global and local spatial autocorrelation, as well as alternative regression models, we investigate the relation between the levels of spatial concentration of the flows themselves (Spanish exports), and the location with regard to the French Border considering distance, travel time and transport costs.

iii. Finally, by means of a gravity equation, we analyse in more detail the structure of the international flows. For this purpose, we decompose the distance measure between the exporting and the importing region in two elements as it is described in Figure 2: from the exporting region to the frontier, and from there to the importing region in the importing country. First, the analysis is based on the whole dataset (with zero values) using a Pseudo-Poisson estimator (PPML) with alternative specifications for the distance variable.

---

6 The use of PPML estimator for gravity equations was proposed by Santos Silva and Tenreyro (2006) for dealing with samples with a larger number of zeros (60% in our case). This estimator also sorts out the Jensen’s inequality (note that
4.1. The Spanish exports to the EU

4.1.1. The shape of the kernel distribution of the outflows

Following some recent articles (Hillberry and Hummels, 2008; Llano-Verduras et al, 2011), we offer here a first view of the distribution of Spanish exports against distance. With this aim, we use a kernel regression to provide a nonparametric estimate of the relationship between the distance from the frontier to the importing region \( \text{dist}_{ij} \) and the intensity of the Spanish exports flows to the regions of the European countries considered corrected by their corresponding GDP’s. Figure 3 plots the kernel distribution of the 15 Spanish regions included in our sample, considering two groups of regions. In the first group we plot the exporting distributions of the border regions with France (Cataluña, Aragón, Navarra, País Vasco) together with others that, as we will see in next sections, also presents interesting features (La Rioja, Comunidad Valenciana, Murcia and Andalucía). The second group includes the rest of the Spanish regions.

<< Figure 3 about here >>

For the first group, the kernel distribution shows a clear concentration of the Spanish international exports in the shortest distance, even after controlling for the size of the trading partners. Conversely, the rest of the regions are grouped in a more heterogeneous group, whose trade appears to be less concentrated in the shortest distance. Three comments are worth mentioning here: i) first, we want to emphasize that the exports are corrected by the GDPs. As a consequence, the corrected intensities of trade of some strong regions will be downgrading if their have a strong specialisation in trading with larger European regions (this could be the case of Cataluña or País Vasco). ii) Also connected with this fact, it is interesting to highlight that when similar analysis has been done (Llano-Verduras et al, 2011), using international and intra-national region-to-country flows, the strong agglomeration of trade in the shortest distance of the international exports remained hidden; iii) Focusing on the kernels of Group 1, it is interesting to make notice that the shapes obtained for País Vasco, Navarra and La Rioja (the ones closer to the south-west French regions) show mono-tonic decreasing slopes from \( \text{dist}_{ij} = 0 \) km, while the ones for Cataluña, C. Valenciana, Murcia and Andalucía (the ones closer to the south-east French regions) show a bell-shape curve. While the former perfectly match with the hypothesis of finding decreasing intensities of trade in the shortest distance, the increasing part of the distribution observed in the later is odd. A possible explanation of this fact is that the largest exports (corrected by their GDPs) to the French south-east regions do not agglomerate in the very first region after crossing the border but in the second.

4.1.2. The spatial concentration of Spanish exports in the importing regions

...
Next, we explore more in detail the higher spatial concentration of the trade flows between some Spanish regions and their European counterparts. In a nutshell, we wonder: 1) if there is a group of Spanish regions with a significant level of spatial concentration in their exports to the main EU regions⁸ 2) and how important is the accessibility (travel times and transport costs) to the frontier with the European core (border with France) for explaining such concentration. In order to answer these questions, we proceed in the following way:

First, by means of the local G-statistic defined by Getis and Ord (1992) (equation 7), we measure a hot-cold analysis for each of the Spanish exporting regions in 2006⁹:

\[
G_i^{ gw} = \frac{\sum_{j=1}^{n} W_{ij} \hat{y}_{ij} - X \sum_{j=1}^{n} W_{ij}}{S \sqrt{n \sum_{j=1}^{n} W_{ij} - \left( \sum_{j=1}^{n} W_{ij} \right)^2}}
\]

\[S = \sqrt{\frac{\sum_{j=1}^{n} \hat{y}_{ij}^2}{n} - (\bar{X})^2} \]

Where:
\[\bar{X} = \frac{\sum_{j=1}^{n} \hat{y}_{ij}}{n} \]

The hot-cold analysis identifies where a variable with either high or low values cluster spatially. This index measures local spatial autocorrelation by looking the value of a variable in a spot within the context of the value of the same variable in the neighbouring locations. In our case, statistically significant hot spots are European regions \((j_i)\) with high values of imports from a specific Spanish exporting region \((i)\), which are surrounded by other European regions \((j_2)\) with high values of imports from the same Spanish exporting region \((i)\). The local sum for that variable and its neighbours is compared proportionally to the sum of the variables in all possible locations: when the local sum is very different from the expected local sum, and that difference is too large to be the result of random chance, a statistically significant z-score results. For statistically significant positive z-scores, the larger the z-score is, the more intense the clustering of high values (hot spot). For statistically significant negative z-scores, the smaller the z-score is, the more intense the clustering of low values (cold spot).

<< Figure 4 about here >>

The results from the G-statistic computed for the exports of each of the Spanish regions using the 2006 data are reported in Figure 4. Note that in this analysis, Portugal was also included, but the export flows are not corrected by the GDPs of the trading partners. Now, thanks to the hot-cold map, it is very easy to interpret previous results, since it capture the spatial concentration of the European main regional clients for each of the 15⁹ Spanish exporting regions. For each one of them, the red collared regions in Europe are the ones with high

---

⁸ In this section, the concept of “spatial concentration of Spanish exports to the EU” is equivalent to say “the spatial clustering of the European importing regions that receive the exports delivered by each Spanish region”. Conversely, the concept of “spatial concentration of Spanish imports from the EU” is equivalent to say “the spatial clustering of the European regions when exporting to a specific Spanish region.

⁹ The analysis avoids using data related to 2007 due to the singularities observed in the trade flows for this year, which may be slightly affected by the first impacts of the current economic downturn.
relative levels of imports, which are located close to other intensive importing regions. In our view, these final results could be easily connected with the two “groups” of regions suggested in Figure 3 (cluster 1 and 2) according to the kernel distribution of trade with the main regions in Europe. According to that figure, and the spatial concentration patterns found in this final section, it seems reasonable to expect larger spatial concentration of flows in the regions that are contiguous to France with the closest regions in Europe.

The spatial concentration of the flows themselves and the shape of the areas coloured in red (high-high concentration schemes) can be described as the swamped area after a dam break event. For some regions (mainly for the Spanish cross-border regions and some regions in the Mediterranean axis), shows high concentration for flows in the first French region after crossing the national border. Moreover, for some of them, the silhouette drown by the G-statistics clearly shows a spatial clustering in the regions located close to the Spanish-French border (accessibility driven exports).

This finding is in line with the results showed by Lafourcade and Paluzie (2011), where border regions in France tend to trade more intensely with the neighbouring countries. Now, thanks to our novel dataset, we have also showed that border regions in Spain are able to trade more intensively with the nearest regions in France. Moreover, we have also show that this geographical advantage is ‘non-dichotomic’, since also non-border regions located close to the two main strategic gateways considered (Girona and Irun), could also show significant spatial concentration patterns. This result suggests that higher spatial concentration levels in exports (spatial clustering of the importing regions in Europe) are associated with ‘geographical advantages’ and better accessibility of some regions to the rest of the continent. However, for those regions with no geographical advantages (current and historically) in terms of accessibility to the European core, we found more dispersed spatial patterns. For these regions, one may expect that exports (at least the ones delivered by road) are less driven by ‘geographical advantages’ and more by other factors such as pure market size (GDP), social and business connections (intra-firm trade) or sectoral complementarity (vertical integration). We expect to shed more light on these issues in further research.

Based on this graphic evidence, we now compute the Global Moran’s Index of Spatial Autocorrelation for the exports of the 15 Spanish regions against the rest of the regions included in our sample (2004-2007). The analytical formula for the Moran’s Index is described by equation [5]:

\[ I_i^m = \frac{n \sum \sum w_{ij} (\hat{T}_{ij} - \bar{T}_i)(\hat{T}_{kj} - \bar{T}_k)}{W \sum (\hat{T}_{kj} - \bar{T}_k)^2} \]

Where \( \hat{T}_{ij} \) is the export value in euros from a Spanish region \( i \) to an EU region \( j \), while \( \hat{T}_{kj} \) is the equivalent to the export to an European region \( j \). Matrix \( W \) is the sum of all the elements in a weight matrix with the inverse distance between all the regions belonging to the European countries considered in the sample (Euclidean distance was used in this case). The elements of this weight matrix are denoted by \( w_{ij} \), which accounts for the distance between each two alternative EU importing regions \( (j_1, j_2) \) that receive exports from
an specific Spanish exporting region (i). By $\bar{T}$ we denote the total average of exports from this specific Spanish exporting region (i) to all the European regions in the sample.

Then, since the goal is measuring the spatial concentration of trade flows and its relation to being a Spanish border region to France or being a Spanish region located near the border, we regress the Moran’s I obtained for each region during the whole period (2004-2007) with respect to different variables that tries to capture this geographical advantage. As said before, thanks to the singular geographical structure of Spain, we have computed three accessibility measures from each Spanish region to the two main corridors thru the Spanish-French border, defined in terms of optimal distance ($dist_{ij}$), optimal travel time and minimum generalized transport costs. Based on these variables, we run the following regressions:

$$Moran' I(z - score)_{i} = \beta_0 + \beta_1 Cont \_French \_Border_i + \varepsilon_i$$ \hspace{1cm} [6]

$$Moran' I(z - score)_{i} = \beta_0 + \beta_1 Prox \_Girona_i + \beta_2 Prox \_Irun_i + \varepsilon_i$$ \hspace{1cm} [7]

The dependent variable is always the z-score obtained when computing the Moran Index for the exports of the 15 Spanish regions against the regions of the countries\(^\text{10}\) considered during the period 2004-2007. Therefore, the models expressed in equations (6-7) try to explain why the Moran Index of spatial autocorrelation is significant for the exports of some regions and not for others. Thus, for each Spanish region we have 4 observations, and therefore 60 observations in total for each specification (4 years*15 regions=60 observations). Equation [6] is used for models M1 and M2. Equation [7] includes two general variables Prox_Girona and Prox_Irun, which will take different values, producing 3 alternative models labelled as M3, M4 and M5. In summary, the specific variables included in each model are:

- **Model 1 (M1):** the two general variables Prox_Girona and Prox_Irun are substituted by a single dummy variable (Cont_French_Border) that takes: value ‘1’ for Aragón and Navarra (regions that are contiguous to France, but suffer from the Pyrenees additional obstacle to trade); value ‘2’ for Cataluña and País Vasco (main entrance to France); ‘0’ otherwise.

- **Model 2 (M2):** include the same dummy of M1, together with the GDP of the Spanish exporting region.

- **Model 3 (M3):** the two general variables Prox_Girona and Prox_Irun are expressed in terms of optimal distance ($dist_{ij}$) to these two enclaves computed by the GIS and the actual transportation network in Spain.

- **Model 4 (M4):** Prox_Girona and Prox_Irun are expressed in terms of minimum travel time computed by the GIS and the actual transportation network in Spain.

- **Model 5 (M5):** Prox_Girona and Prox_Irun are expressed in terms of minimum generalized transport cost to these two spots computed by the GIS and the actual transportation network in Spain.

<< Table 2 about here >>

\(^{10}\) Since the goal is measuring the “geographical advantage” of being located to the French border for exporting to the “main continental Europe”, Portugal is not considered in this analysis.
Table 2 shows the results obtained for these 5 specifications. M1 and M2 show how the significant level of the spatial autocorrelation of the Spanish regional exports (spatial concentration of the importing regions in the countries considered) is positively correlated with being a border-region with France, as well as the level of ones GDP. M3, M4 and M5 also show that not just being a border region (dichotomic measure of the geographical advantage regarding trading with the rest of Europe) but also being close to the two main gateways (clearer for the case of Girona than for Irun) also plays a role on explaining the spatial concentration of the Spanish regional clients in Europe.

4.1.3. Decomposing the effect of internal and external distance on exports.

In this section the gravity equation is used for analysing in more detail the heterogeneous and non-linear relation between the international trade flows and the distance travelled by the trucks within or outside Spain when exporting to its European partners. For this purpose, we decompose the distance measure between the exporting and importing region into 2 elements, as it is described in Figure 2: on the one hand the distance from the exporting region to the most convenient entrance through the French-Spanish border (Irun versus Girona); on the other, the distance from these points in the frontier to the final destination in the importing country. This analysis is conducted for the aggregate flows, using alternative specifications and estimation procedures. Portugal and Andorra are excluded. The first one is defined by equations [8], [9] and [10]:

\[ T_{ijt} = \beta_0 + \beta_1 \ln Y_i + \beta_2 \ln Y_j + \beta_3 \ln \text{dist}_{ij} + \beta_4 \ln \text{GDP}_i + \beta_5 \ln \text{GDP}_j + \gamma_t + \epsilon_{ij} \]  
\[ \frac{T_{ijt}}{Y_i Y_j} = \beta_0 + \beta_1 \ln \text{dist}_{ij} + \beta_2 \ln \text{GDP}_i + \beta_3 \ln \text{GDP}_j + \gamma_t + \epsilon_{ij} \]  
\[ \frac{T_{ijt}}{Y_i Y_j} = \beta_0 + \beta_1 \ln \text{dist}_{ij} + \beta_2 \ln \text{GDP}_i + \beta_3 \ln \text{GDP}_j + \gamma_t + \epsilon_{ij} \]

where \( T_{ijt} \) is the flow from the region \( i \) in country \( e \) to the region \( j \) in country \( u \) at year \( t \). Note that if \( e \neq u \), such term captures interregional flows between Spain and another European country in the sample (Spanish exports if \( e=\text{Spain} \)). The variables \( \ln Y_i \) and \( \ln Y_j \) represent, respectively, the logarithm of the nominal gross domestic product (GDP) of the exporting and the importing region. Equations [9] and [10] are expressed in terms of the ratio between the bilateral trade flows and the GDPs of the trading regions (both in levels).

As it was explained in the previous section, the variable \( \text{dist}_{ij} \) (actual total distance traveled by the trucks) between the Spanish exporting region \( i \) and the European importing region \( j \), has been split in two parts: a) \( \text{dist}_{ij} \), which captures the distance from the exporting region \( i \) to the most convenient gateway to the destination market through the French-Spanish border (Girona or Irun); b) then, as explained before, \( \text{dist}_{ij} \), is obtained as a residual \( \text{dist}_{ij} = \text{dist}_{ij} - \text{dist}_{ij} \), and captures the distance from the most convenient gateway through the French-Spanish border (Girona or Irun) to the final destination in the importing country \( j \).
In addition, a new variable \( \text{dist}_{ij}^2 \) is included. Note that, in contrast to other previous works (Hillberry and Hummels, 2008; Llano-Verduras et al, 2011; Gallego and Llano 2014, 2015) these variable makes reference to the square value of the distance between the frontier and the destination region \( \text{dist}_{ij} \), and not to the square of the total distance \( \text{dist}_i \). By doing so, we expect to capture the non-linear relation between trade flows and the distance traveled in the destination country. Thus, it will be capturing the agglomeration of inflows in the shortest distance just after crossing the border, an effect that was observed for some regions in the kernel regressions (Figure 3) and the hot-cold analysis (Figure 4). The interpretation of the three distance variables (introduced in levels) included is straightforward: i) a negative and direct effect of \( \text{dist}_{ij} \) and \( \text{dist}_{ij} \) on trade is expected; ii) in addition, for those regions whose export flows are highly concentrated in the regions that are close to the French border, a positive sign is expected for the square of the distance from the frontier to the importing region \( \text{dist}_{ij}^2 \). Therefore, according to the two paradigmatic spatial patterns of trade described in Figure 2, we expect to obtain a positive and significant coefficient for the regions with ‘trade-cost-driven-trade’, and a non-significant or even negative coefficient for the ‘market-size-driven-trade’.

In addition, we use also a set of alternative dummy variables to control for the effect of an expected more intense trade in the French regions that are closer to the border-main entrees in Girona and Irun. We remark to this regard what was said before regarding the presence of strong cultural, geographical and historical ties between these border regions:

- **Aquit**: is a dummy variable that takes value 1 when a Spanish export from region \( i \) has a final destination \( j \) in Aquitania, which is the French border region contiguous to País Vasco in Spain (where Irun is). It takes value 0 otherwise.

- **C_Aquit**: is a dummy variable that takes value 1 when a Spanish exporting region \( i \) trades with any of the contiguous regions of Aquitania in France, and value 0 otherwise.

- **Lang_R**: is a dummy variable that takes value 1 when a Spanish exporting region \( i \) trades with Languedoc-Roussillon, which is the French border region contiguous to Cataluña in Spain (where Girona is). It takes value 0 otherwise.

- **C_Lang_R**: is a dummy variable that takes value 1 when the Spanish exports goes from \( i \) to \( j \), being \( j \) a contiguous region of Languedoc-Roussillon in France, and value 0 otherwise.

Finally, the terms \( \mu_i \) and \( \mu_j \) correspond, respectively, to the multilateral resistance fixed effects for each origin and destination region by year. The variable \( \gamma_t \) is the time fixed effect.

Table 4 reports the results obtained for 8 alternative specifications using the whole sample of the Spanish exports to the European regions considered (omitting Andorra and Portugal).

<< Table 4 about here >>

From a large number of alternative specifications (all of them available upon request), Table 4 reports the main ones able to capture the spatial concentration of Spanish exports in the French regions that are closer to the border. All of them are run using the whole sample (with 60% of zero flows) and the PPML estimator. The
ones in column (1) are based in equation [8] and use the Spanish exports as dependent variable and just includes the time fixed effects. The coefficients for all the variables included in (1) are significant and have the expected signs. The coefficients for the log of the GDP’s are close to the unity. This specification shows how Spanish exports decrease with the internal distance (distfj) as well as with the external one (distf). The quadratic term (distfj²) obtains a positive and highly significant coefficient (0.902***), which indicates a more than proportional agglomeration of trade in the shortest distance after crossing the French border.

In coherence with the literature (Anderson and van Wincoop, 2003), the rest of the specifications include fixed effects for the origin and the destination regions, and uses the corrected trade export flows as dependent variable (\(T_{ij}^{out}/Y_tY_R\)). As a consequence, the distance from the exporting region to the frontier (distf), which is an invariant scalar for each exporting Spanish region, will be subsumed by the origin-fixed effect, becoming non-significant. For this reason, the later drops from all the subsequent specifications.

Like in (1), the specifications (3, 6, and 7) include the variables distfj and distfj² in levels. The results in column (3) are equivalent to the ones in (1), showing that the strong agglomeration of trade in the shortest distance after the frontier is also captured when the dependent variable is expressed as a ratio and the multilateral terms are included. In specification (6) the dummy variable for the French regions nearest to Girona (Cataluña) are included. Although the coefficients for the distance variables remain almost invariant with respect to (3), the coefficient for Lang_R is positive and significant, while the one for C_Lang_R is negative and significant.

Complementary, in specification (7) the dummy variable for the French regions near to the two main entrée doors (Irun and Girona) are included. As a consequence, the negative value for the coefficient of distfj drops from -5 to -3 and distfj² becomes non-significant. Now the effect of concentration of trade in the French border regions is captured by the two positive and significant coefficients of Aquit and Lang_R. Specification (4) focus on the effect of the flows entering thru Girona (Cataluña), finding positive and significant effects for Lang_R and negative for their neighbours C.Lang_R. Similarly, specification (5) considers the case of the French border regions in the south-west, finding non-significant effects. In specification (8) all the dummies are included at the same time, showing no significant results for the individual dummies, but obtaining a drop in the coefficient for the log of distance.

Now, we repeat the analysis using different sets of sub-samples, which are defined considering different groups of Spanish regions. The regions have been grouped according to the kernels distributions and the hot-cold spatial analysis: we define two sets of groups with the aim of identifying regions whose exports are more likely to experience geographical advantages with regards to the continental Europe, and therefore, may show higher spatial agglomeration patterns. Table 5 reports the results for some alternative specifications considering the possibility larger agglomeration patterns in the exports delivered by the Spanish border regions with France (Pais Vasco, Navarra, Aragón and Cataluña).

<< Table 5 about here >>
Table 5 is split in two panels. In the first one (Group 1: columns 1 to 5) the results for the Spanish border regions to France are included. In the second (Group 2: columns 5 to 10) the same specifications are reported for the rest of the country. Results in column (3) and (8) shows that exports from both groups have negative and significant coefficients for \( \text{dist}_f \) and positive and significant coefficients for \( \text{dist}_f^2 \). However, they also show that the Spanish-border-regions have a larger concentration than the average (-6.483*** vs -5.052*** for \( \text{dist}_f \); 2.175*** compared to 1.948** for \( \text{dist}_f^2 \)). This result may be odd if ones do not consider the import-concentration hypothesis described before, since there will be no reason why the most accessible regions should have to show higher concentration in the shortest distance rather than in the largest ones, which in this case are the furthest ones (North of France, Belgium and Germany).

Like before, the rest of the specifications in Table 5 try to control for the same effect by means of adding specific dummies. For the Spanish border regions non-significant results are obtained for such dummies, while for the non-border Spanish region positive and significant results are obtained with regards to Languedoc-Roussillon (0.804*** in specification 7, and 0.612***, in 10) and negative and significant ones with regards to Languedoc-Roussillon’s neighbors (-0.954*** in specification 7, and -1.044***, in 10). In our view this can be interpreted as an indication that the concentration of exports from the Spanish border regions goes further away than the pure contiguous French regions, while many non-border regions show an unexpected strong concentration of exports in the very first French region just after crossing the border (see Figure 4). Although Gallego and Llano (2012) analysed the current dataset in parallel to the official trade data, and discarded severe problems of transit flows contaminating our shipments, more research is needed in order to explain this idiosyncratic larger demand of Spanish exports arriving to Languedoc-Roussillon (and Aquitania to a lower degree) from several Spanish regions. This could be a consequence of complex logistic practices (multi-modal transport connections or intra-firm storage facilities), but also an effect of the strong historical ties that this specific regions has kept with Spain throughout History (see section 4.1.).

Although the results obtained partially support our hypothesis of larger agglomerations of trade in the nearest regions to the border, the fact of finding positive and significant effects of this kind not just for the Spanish border regions but for all, encourage us to make a robustness analysis looking for complementary explanations. With this aim Table 6 shows the results for an alternative geographical clustering of the Spanish regions, which in fact could also be linked with the idea of geographical advantage, but also to sectoral specialization. Now, Spain will be divided in Group 3 (the Mediterranean regions: Cataluña, Comunidad Valenciana, Murcia and Andalucía) and Group 4 (the rest of the country). Note that if one consider the Pyrenees as a real physical barrier for interacting with the continental Europe, a longitudinal divide of the peninsula is as justifiable as the north-south division considered before.

<< Table 6 about here >>

Based on the results reported in Table 6 the second grouping strategy is also very informative. The exports delivered by regions in Group 3 (the Mediterranean regions) to the French regions located in the south-west \((\text{Aquit} \text{ and } \text{C_Aquit})\), are significantly below the average (column 1), while the ones to \( \text{Lang}_R \) are
significantly above the average (column 2). In fact, as it is reported by specification (4), once that the exports from the Mediterranean regions to Aquit and C_Aquit are controlled for, the results obtained for dist$ij$ and the quadratic term (dist$ij^2$) indicates a more than proportional concentration of trade in the French regions located close to the south-east entrée door through Girona. Conversely, when we focus on the second panel of the table (Group 4: rest of the country excluding the Mediterranean regions), the opposite is observed: now the exports delivered to the French regions located in the south-east (Lang_R and C.Lang_R), are negative and significant (column 7), and the inclusion of these two dummies together to dist$ij$ and dist$ij^2$ (column 10) leads to a more than proportional concentration of trade in the French regions located close to the French south-west gateway (Irun). We interpret this result as a sign of a kind of geographical specialization of the Spanish regions when exporting to Europe (France specifically). Interestingly, this division do not follow the criteria suggested by the literature between ‘border-regions versus non-border regions’ (north-south divide in our case), but an alternative one following an east-west axis. Such result, far from contradicting our hypothesis on the existence of a larger spatial concentration of imports for certain regions (the ones with geographical advantages) it will enrich it introducing the idea of river basins: if Spanish exports behave as water, and have to pass through the two existing extreme funnels (Irun and Girona) for overpassing the Pyrenees physical barrier, the spatial agglomeration of the inflows entering in France would depend more on the proximity to each of this gateways, than just on the condition of being or not a border-region. Such result can be expressed in terms of a geographical specialization of the eastern-western regions with regards to the closest French regions, a fact that can be described in our natural metaphor by considering the effect of an obstacle in the course of the fluid, or by considering different slopes of the surfaces receiving the flows in the importing country.

4.2. A product specific analysis. [to be developed]

In this section we include a sector specific analysis.

The aim is to test to what extent the value/volume ratio for each specific product $g$ exported abroad also drives the level of agglomeration of imports in the destination country.

<< Figure 4 about here >>

Figure 4 may include:

- Panel A) a scatterplot with 3 dimensions: 1) dist$_{ij}$ (y axis); 2) dist$_{ij}$ (x axis); 3) size of the ball: 
  \[ \ln(\text{value}_{ijt}^g) \text{ or just } T_{ijt}^g \text{ (average of the period, or just 1 year, 2006)} \]

- Panel B) a scatterplot with 3 dimensions: 1) \( \frac{\text{value}_{ijt}^g}{\text{volume}_{ijt}^g} \) (y axis); 2) dist$_{ij}$ (x axis); 3) size of the ball: 
  \[ \text{dist}_ij^2 \text{. Average of the period, or just 1 year, 2006} \]

Finally, we test the following specifications:

\[
\frac{\text{value}_{ijt}^g}{\text{volume}_{ijt}^g} = \beta_0 + \beta_1 \text{dist}_{ij} + \beta_2 \text{dist}_{fj} + \beta_3 \text{dist}_{ij}^2 + \beta_4 \text{Aquit} + \beta_5 \text{C_Aquit} + \beta_6 \text{Lang}_R + \beta_7 \text{C_Lang}_R + \mu_{jt} + \mu_{jt} + \mu_g + \epsilon_{ijt} \quad [10]
\]
Equation [10] will allow analyzing the heterogeneity of product specific flows with respect to the variables and specifications already tested for the aggregate flows:

\[ AT^g_t = \beta_0 + \beta_1 \text{dist}_{ij} + \beta_2 \frac{T(\text{value})^g_{ijt}}{T(\text{volume})^g_{ijt}} + \mu_g + \varepsilon_{ij} \]  

Then, we define \( AT^g_t = W \times T(\text{value})^g_{ijt} \) as the ‘aggregated trade’ of product \( g \) in year \( t \), computed as the weighted average of the flows arriving to each foreign importing region \( j \), weighted by the distance to the gateway (dist\(_f\)) of entrance when delivered by each Spanish exporting region \( i \).

Our intuition is that the lower the \( \frac{T(\text{value})^g_{ijt}}{T(\text{volume})^g_{ijt}} \) and the \( \text{dist}_{if} \), the larger the level of concentration of flows at the other side of the French border. For products \( g \) fulfilling this hypothesis, we will be confirming the presence of “transport-cost driven trade”. Additional variables such as the “herfindal” index for each importing region can also be included for each \( g-j \) (capacity of \( j \) of exporting products \( g \) to Spain or to the world)

<< Table 7 about here >>

5. Conclusions

When considering a world of heterogeneous space and asymmetric regions, some areas -like border regions- might benefit from geographical advantage when exporting to foreign markets. This fact raises different questions about the internal allocation of the exporting firms as well as the spatial pattern of the flows that they will generate. To date, the literature has concentrated mainly on analysing the internal effects of openness to trade, by measuring to what extent the liberalisation process of goods and factors mobility triggers the internal agglomeration of firms and jobs close to the regions with geographical advantages (i.e. border regions). By contrast, little attention has been devoted to the destination side of the issue, that is, the spatial concentration of the flows themselves when arriving to the importing countries.

With the aim of filling this gap, in this paper we have explored the presence of different spatial pattern in the concentration of the inflows in the importing regions depending on the location advantage of the exporting region (located close or further away from the national border). By doing so, we wonder if the spatial pattern of border regions are mainly explained by this geographical advantages (accessibility).

This approach was possible thanks to a new dataset that considers interregional trade flows shipped by Spanish trucks between the Spanish regions to the regions of the 8 main European partner countries. This dataset also considers the distance associated with the shipment of the flows, which in our case has been split in two parts, the distance from the exporting region to the frontier (internal geography) and from there to the importing region in the foreign country. Departing from this novel dataset, a number of analyses are developed with the aim of identifying different spatial patterns of the Spanish international trade, with the region-to-region specific breakdown. Through these analysis, we were able to show that those regions located close to the Spanish-French border have a statistically significant propensity to trade more with the regions located closer to the border, even when controlling for their GDPs. Furthermore, unexpectedly, we have also identify an interesting
west-east specialisation of the Spanish regions when exporting to Europe by truck, which is mainly explained by the presence of a strong natural barrier to trade (Pyrenees) with just two main gateways located more than 600km one from the other.

In future research, we expect to further investigate to what extent these alternative spatial patterns are explained more by current advantages in transport cost, or by a historical accumulation of exporting firms willing to allocate within the country but as close as possible to the border. For doing so, we will explore the industry specific dimension of this dataset, in search of possible selection bias of the firms located close to the border with regards to products of low transportability or low levels of competence at the other side of the border. Furthermore, alternative drivers such as social and business networks would be considered for explaining the performance of some non-border regions, whose trade flows are more fond of traveling larger distances.

6. References


Table 1. Spatial concentration of regional exports and proximity to the border. OLS regression. Dependent variables: Moran Index z-score statistic, obtained for the exports in monetary flows, for each Spanish regions against the other EU regions. (Portugal excluded). Pooled data for 2004-2007.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
</tr>
</thead>
<tbody>
<tr>
<td>cont_french_border</td>
<td>1.140***</td>
<td>1.016***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.178)</td>
<td>(0.178)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gdp_pps</td>
<td>0**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dist_Girona</td>
<td>-0.00288***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000626)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dist_Irun</td>
<td>0.00125*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000649)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_Girona</td>
<td>-0.00315***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000700)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_Irun</td>
<td>0.00110</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000749)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cost_Girona</td>
<td>-6.61e-05***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.59e-05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cost_Irun</td>
<td>1.25e-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.88e-05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.081***</td>
<td>0.794***</td>
<td>2.910***</td>
<td>2.886***</td>
<td>2.830***</td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.181)</td>
<td>(0.416)</td>
<td>(0.456)</td>
<td>(0.482)</td>
</tr>
<tr>
<td>Observations</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.413</td>
<td>0.471</td>
<td>0.270</td>
<td>0.264</td>
<td>0.233</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Table 2. Non-linear response of trade to the distance at both sides of the French border.
Complete sample. PPML. Region-to-region Spanish exports to Europe. (Andorra & Portugal excluded).

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$II_i$</td>
<td>0.870*** (0.0671)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$II_j$</td>
<td>0.891*** (0.0825)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ID_f$</td>
<td>-0.988*** (0.322)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_f$</td>
<td>-0.945*** (0.220)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_f$</td>
<td>-3.645*** (0.641)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_f^2$</td>
<td>0.902*** (0.312)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Aquit.$</td>
<td>1.109** (0.532)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_Aquit.$</td>
<td>-0.453 (0.475)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Lang_R$</td>
<td>1.087*** (0.305)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_Lang_R$</td>
<td>-0.603*** (0.151)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.909 (1.205)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| FE by year | YES |
| FE orig. x year | NO | YES | YES | YES | YES | YES | YES | YES |
| FE dest. x year | NO | YES | YES | YES | YES | YES | YES | YES |
| Observations | 7,515 | 7,515 | 7,515 | 7,515 | 7,515 | 7,515 | 7,515 | 7,515 |
| R-squared | 0.211 | 0.211 | 0.211 | 0.211 | 0.211 | 0.211 | 0.211 | 0.211 |

Standard errors in parentheses GDP is measured in term of Purchasing Power Standard. . *** p<0.01, ** p<0.05, * p<0.1
Table 3. Non-linear response of trade by Groups: “Spanish border regions” vs “Rest of Spain”. PPML.
Region-to-region Spanish exports to 6 EU countries. (Andorra and Portugal excluded), 2004-07.

<table>
<thead>
<tr>
<th>Variab.</th>
<th>Group 1: border regions with France†</th>
<th>Group 2: rest of the Spanish regions (islands excluded)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>lD_y</td>
<td>-1.982***</td>
<td>-1.912***</td>
</tr>
<tr>
<td>Aquit.</td>
<td>-0.0824</td>
<td>0.408</td>
</tr>
<tr>
<td>C_Aquit.</td>
<td>-0.3977</td>
<td>-0.438</td>
</tr>
<tr>
<td>Lang_R.</td>
<td>0.222</td>
<td>-0.0506</td>
</tr>
<tr>
<td>C_LangR.</td>
<td>-0.2484</td>
<td>-0.452*</td>
</tr>
<tr>
<td>D_y</td>
<td>-6.483***</td>
<td>-5.596***</td>
</tr>
<tr>
<td>D_y²</td>
<td>2.175***</td>
<td>1.718***</td>
</tr>
<tr>
<td>Const</td>
<td>-4.602***</td>
<td>-4.614***</td>
</tr>
<tr>
<td>Observ.</td>
<td>1,716</td>
<td>1,716</td>
</tr>
<tr>
<td>R²</td>
<td>0.570</td>
<td>0.570</td>
</tr>
<tr>
<td>µ_k</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>µ_j</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Standard errors in parentheses GDP is measured in terms of Purchasing Power Standard. . *** p<0.01, ** p<0.05, * p<0.1
µ_k: FE by origin & year; µ_j: FE by origin & year; †Group 1: border regions with France: País Vasco, Navarra, Aragón and Cataluña.

Table 4. Non-linear response of trade by Groups: “Mediterranean Axis” vs “Rest of Spain”. PPML.

<table>
<thead>
<tr>
<th>Variab.</th>
<th>Group 3: Mediterranean Axis †</th>
<th>Group 4: rest of the Spanish regions (islands excluded)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>lD_y</td>
<td>-1.911***</td>
<td>-0.730***</td>
</tr>
<tr>
<td>Aquit.</td>
<td>-2.231***</td>
<td>-2.022***</td>
</tr>
<tr>
<td>C_Aquit.</td>
<td>-1.226**</td>
<td>-1.355**</td>
</tr>
<tr>
<td>Lang_R.</td>
<td>2.005***</td>
<td>2.106***</td>
</tr>
<tr>
<td>C_LangR.</td>
<td>0.0728</td>
<td>0.0681</td>
</tr>
<tr>
<td>D_y</td>
<td>-2.892**</td>
<td>-6.508***</td>
</tr>
<tr>
<td>D_y²</td>
<td>0.733</td>
<td>2.254***</td>
</tr>
<tr>
<td>Const</td>
<td>-4.492***</td>
<td>-4.654***</td>
</tr>
<tr>
<td>Observ.</td>
<td>1,924</td>
<td>1,924</td>
</tr>
<tr>
<td>R²</td>
<td>0.570</td>
<td>0.570</td>
</tr>
<tr>
<td>µ_k</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>µ_j</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Standard errors in parentheses GDP is measured in terms of Purchasing Power Standard. . *** p<0.01, ** p<0.05, * p<0.1
µ_k: FE by origin & year; µ_j: FE by origin & year; †Group 3: Mediterranean Spanish regions (Cataluña, Valencia, Murcia, Andalucía)
Figure 1. Scheme describing two alternative patterns of trade.

<table>
<thead>
<tr>
<th>A) Trade cost driven trade</th>
<th>B) Market size driven trade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pais Vasco</td>
<td>Paris</td>
</tr>
<tr>
<td>Cataluña</td>
<td>Rome</td>
</tr>
<tr>
<td></td>
<td>Madrid</td>
</tr>
</tbody>
</table>

FIGURES
Figure 2: Kernel regression: Inter-national exports relative to the GDP on $d_{ij}$ by region. $T_{ij}(Y_i, Y_j)$.

Group 1: Spanish exporting regions:
País Vasco, Navarra, La Rioja, Aragón, Cataluña, Com.Valenciana, Murcia and Andalucía.

Group 2. Spanish exporting regions:
Galicia, Asturias, Cantabria, Madrid, Castilla y León, Castilla-La Mancha, Extremadura
Figure 3: The Hot-Cold graph based on Getis&Ord Spatial Autocorrelation Index. Exports from the Spanish regions (star) to the main EU regions, 2006.
Table 5: Countries and regions included in the sample. 2004 to 2007.

<table>
<thead>
<tr>
<th>Countries</th>
<th>NUTS 0</th>
<th>Spanish Regions</th>
<th>NUTS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andorra</td>
<td>AD</td>
<td>Andalucía</td>
<td>ES61</td>
</tr>
<tr>
<td>Belgium</td>
<td>BE</td>
<td>Aragón</td>
<td>ES24</td>
</tr>
<tr>
<td>Germany</td>
<td>DE</td>
<td>Asturias</td>
<td>ES12</td>
</tr>
<tr>
<td>Spain</td>
<td>ES</td>
<td>Baleares</td>
<td>ES53</td>
</tr>
<tr>
<td>France</td>
<td>FR</td>
<td>Canarias</td>
<td>ES70</td>
</tr>
<tr>
<td>Italy</td>
<td>IT</td>
<td>Cantabria</td>
<td>ES13</td>
</tr>
<tr>
<td>Netherlands</td>
<td>NL</td>
<td>Castilla y León</td>
<td>ES41</td>
</tr>
<tr>
<td>Portugal</td>
<td>PT</td>
<td>Castilla La Mancha</td>
<td>ES42</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>UK</td>
<td>Cataluña</td>
<td>ES51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comunidad Valenciana</td>
<td>ES52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extremadura</td>
<td>ES43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Galicia</td>
<td>ES11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comunidad de Madrid</td>
<td>ES30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Región de Murcia</td>
<td>ES62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Navarra</td>
<td>ES22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>País Vasco</td>
<td>ES21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>La Rioja</td>
<td>ES23</td>
</tr>
</tbody>
</table>