Islands in trade: disentangling distance from border effects

José L. Groizard
<joseluis.groizard@uib.es>

Helena Marques
<helena.ferreira-marques@uib.es>

María Santana
<maria.santana@uib.es>

Department of Applied Economics
University of the Balearic Islands

Abstract

There is a well-established literature on border effects covering trade between regions separated by a land border; however that literature has not so far considered the case of regions separated by a sea border. Whilst the former is typically studied as a political border that affects adjacent regions belonging to different countries and can be reduced by free trade agreements, the latter is a geographical border that affects regions within the same country and cannot be reduced in a similar way. Both types of borders produce similar effects upon trade. However, in the case of the sea border there are specific transport costs to be considered due to the existence of a time barrier, the need to combine different transport modes – usually truck plus ship or alternatively truck plus air transport – or to pay fees and taxes for the use of public infrastructures such as ports and airports. The empirical strategy used to estimate the "island effect" proceeds in two steps. First an augmented gravity model is estimated for mainland and island regions; then a Blinder-Oaxaca decomposition is applied to the gravity estimation results in order to disentangle the distance and border effects for those regions, net of all other factors controlled for in the gravity estimations. Results show that island regions are at a substantial disadvantage compared to continental regions, which is due to the higher and non-linear effect of distance coefficients.

Keywords: F15, C23
JEL codes: Gravity equation; Border effects; Panel data; Spain; Regional trade

1. Introduction

There is a well-established literature on border effects covering trade between regions separated by a land border;[1] however that literature has not so far considered the case of regions separated by a sea border. This is an important distinction because, whilst the former is typically a political border that affects adjacent regions belonging to different countries and can be reduced by free trade agreements,[2] the latter is a geographical border that affects regions within the same country and cannot be reduced in a similar way. Both types of borders produce similar effects upon trade. However, in the case of the sea border there are specific transport costs to be considered due to the existence of a time barrier, the need to combine different transport modes – usually truck plus ship or alternatively truck plus air transport – or to pay fees and taxes for the use of public infrastructures such as ports and airports.[3] All those additional costs are different from the traditional frictions considered in the literature (e.g. tariffs, language barriers, etc.) and require a proper specific treatment.

In this paper, we measure the trade effects of the existence of a sea border for the case of Spain, a country that includes two island regions: Balearic Islands and Canary Islands (see Figure 1). The structure of the trade cost of a typical shipment between the Balearic Islands or Canary Islands and any mainland region is different than that between two mainland regions. The first one involves typically the use of two transport modes, by road and by ship. The second one involves only the use of road transport.[4] Although most studies use as a proxy for transport costs the distance between origin and destination the nature of those costs depend crucially on the transport mode used in the shipment. Ground transportation costs are somehow observable while maritime transportation costs are not, especially for interregional shipments and this makes difficult to ascertain the relative costs of both modes.[5] However, the literature provides a guide to consider some major features that suppose a penalty for the maritime transportation cost, such as the insurance cost, the especial conditions required for the type of goods shipped (e.g. refrigerated transport, fuel, etc.), small economies of scale at the port-level and at the cargo-ship-level, legal regulations and the existence of anticompetitive practices.[6] We call all those extra costs of trade Island-specific costs, which do not exist when trade takes place between regions located in the Iberian Peninsula, and these costs taken together constitute the "island effect".[7]

---

[1] See, among many others, Anderson and van Wincoop (2003), Chen (2004), Evans (2003), Head and Mayer (2000), McCallum (1995).

[2] US and Canadian regions or European Union countries are the most studied cases (see footnote 1).

[3] Ground transportation also uses public transport infrastructure (road usually), however the Government does not tax or charge specific fees for that use and the funding of maintenance costs depends on direct and indirect taxes and heavy subsidies.

[4] As we will see in Section 3, the truck is the most common means of transport in the delivery of goods that are produced in the peninsula, while the combination of truck-ferry is the preferred option for trade between the islands and the mainland.

[5] Official statistics contain detailed information to measure and identify the various components of the costs of road transport in Spain, however, unfortunately, does not reflect any statistics on the costs of transportation by sea.

[6] See for example, Clark et al. (2004) and Marquez-Ramos et al. (2007).

[7] Some other extra costs arise because the maritime transport mode is slower and the time loss is sometime unexpected when the route has to be cancelled due to adverse sea conditions. This time delay may be much more costly for a firm than just the other costs to be paid, but it is hard to place a monetary value on it. This type of

Figure 1: Regions of Spain

NOTE: Kindly and freely available from http://www.map-of-spain.co.uk/

These extra costs are likely to be relevant when trade is interregional and not international due to the enormous savings that exist when the scale of operations is large.[8] Therefore, in this paper we focus on interregional trade. The empirical strategy used to estimate the "island effect" proceeds in two steps. First an augmented gravity model that includes all types of trade costs incurred by island regions within their country is estimated for mainland and island regions; then a Blinder-Oaxaca decomposition is applied to the gravity estimation results in order to disentangle the distance and border effects for those regions, net of all other factors controlled for in the gravity estimations.

On the one hand, the gravity model has been for a long time the most widely used empirical model of International Economics research.[9] The most basic formulation of the gravity model consists of explaining bilateral flows as a direct function of the two partners' economic size (measured in terms of GDP, GDP per capita and/or population) and as an inverse function of the distance between them (Anderson, 2011). In the current paper we propose an extension of the approach taken in Spies and Marques (2009) to incorporate a simplified version of the trade cost function of Novy (2013) and Feenstra and Romalis (2014), who consider two types of trade costs in the context of international trade: the usual "iceberg" variable cost, which depends on distance, and a specific trade cost, which exists only for some trade partners (in our case, islands). Thus we distinguish the effect of distance (the typical gravity model iceberg

island-specific cost does not exist between, for example, Barcelona (Catalonia) and Vigo (Galicia) because transport can be carried out by road in a timely manner, despite the distance.

[8] In long distance trade containerized transport is intensively used and the scale of operation allows benefiting from more efficient newer generations of cargo-ships. Moreover, the large scale of operations leads to a higher number of lines connecting ports and, thus, increasing the competitive pressure in the transport market and, as a result, lowering the cost of transport services.

[9] These models have been applied to trade (Marques and Metcalf, 2005, 2006; Papazoglou et al., 2006; Armstrong 2007; Spies and Marques, 2009; Marques, 2011), migration (Gil-Pareja et al. 2006; Marques, 2010), FDI (Head and Ries 2008), and tourism (Eilat and Einav, 2004; Gil-Pareja et al., 2007; Santana et al, 2010; Fourie and Santana 2013; Rosselló and Santana-Gallego 2014). Moreover, this specification has also been used both in the international and regional context (see, for the case of Spain, Sansó et al., 1990, Sanz, 2000 and Gil-Pareja et al., 2005).

cost) from the effect of the existence of a sea border (the partner-specific trade cost). This is an important issue for island economies because distance and border effects play a different role for those economies.

On the other hand, the Blinder-Oaxaca approach allows us to decompose in two components why islands trade less than mainland regions. A first explanation is that they have different characteristics (lower size or greater distance to other regions). A second explanation is that they face different "prices" associated to those characteristics (i.e. a great penalty to transport merchandise with origin or destination in the islands). Based on the gravity estimates, we isolate those effects by generating two types of counterfactuals. The first one sets how much trade the two Island regions would carry out if they were affected by the same coefficients (i.e. same transportation costs) while keeping their own characteristics (i.e. distance differs as well as the other features). This way we isolate the effect of characteristics. The second counterfactual sets how much trade there would be if the characteristics between the two types of regions were the *same* (i.e. same distance, etc.) but the coefficients were *different* (i.e. trade costs differ for islands and non-island regions). This way we isolate the effect of the premium or penalty that regions face (i.e. discrimination in transport costs).

Results show that island regions are at a substantial disadvantage compared to continental regions, but their trade disadvantage is due to a greater extent to the higher and non-linear effect of distance. We find that, for both type of regions, trade lowers when distance increases, following an asymmetric U-shape functional form, that is, trade is lowered more at intermediate and large distances than at shorter distances. When we compare the effect of distance coefficients for the samples of mainland and island regions we find that coefficients are much higher for islands than for mainland regions. These results lend support to our hypothesis that there are specific costs when trade involves dealing with island regions that are different from other frictions related to trade. The bottom line is that such trade cost frictions are responsible for a large gap of trade between mainland and island regions. For instance, the quantitative results show that lower relative trade costs for mainland regions generate an export gap from 1 to 1.6 log points and an import gap from 0.6 to 1 log points greater for mainland regions. These effects seem to be relevant given that mainland regions export 3.1 log points and import one-log point more than island regions.

In what follows, section 2 explains the two-step empirical strategy used to estimate the "island effect". Section 3 describes the data sources and the main features of those regions' international and interregional trade structure. Section 4 presents the empirical results. Section 5 describes all the robustness checks carried out and section 6 concludes.

2. Empirical strategy

2.1. Augmented gravity model

Assuming identical, homothetic Constant Elasticity of Substitution (CES) preferences, region i's aggregate total value of imports from region j in year t (M_{ijt}) can be expressed as:

$$M_{ijt} = N_{jt} Y_{it} \left(\frac{p_{ijt}}{P_{it}}\right)^{1-\sigma} \tag{1}$$

with N_{jt} representing the number of products sold by region j, Y_{it} being region i's nominal expenditure (measured by its GDP); $\frac{p_{ijt}}{P_{it}}$ is the relative price determining the share of region i's GDP allocated to purchasing imports from region j, with P_{it} being region i's price index for all import-competing goods (whether produced in the region or in third regions) and p_{ijt} standing for the price at destination (in region i); finally, $\sigma > 1$ is the elasticity of substitution between goods originating from the two trading regions i and j.

Assuming the existence of both partner-specific (T_{ij}) and "iceberg" variable (τ_{ij}) bilateral trade costs, the price at destination (p_{ijt}) is defined as:

$$p_{ijt} = \tau_{ij} P_{jt} + T_{ij} \tag{2}$$

where P_{jt} is region *j*'s producer price index.

Substituting (2) into (1) yields:

$$M_{ijt} = N_{jt} Y_{it} \left(\frac{\tau_{ij} P_{jt}}{P_{it}} + \frac{T_{ij}}{P_{it}}\right)^{1-\sigma} \tag{3}$$

Under general equilibrium, region *j*'s producer price must adjust such that the market clearing condition is satisfied. Assuming instantaneous adjustment, which seems fairly plausible at the regional level, we have that:

$$Y_{jt} = \sum_{i=1}^{I} M_{ijt} \tag{4}$$

Substituting the import demand equation (3) into the market clearing condition (4), we can solve for N_{jt} as follows:

$$N_{jt} = \frac{Y_{jt}}{\sum_{i=1}^{I} Y_{it} \left(\frac{\tau_{ij} P_{jt}}{P_{it}} + \frac{T_{ij}}{P_{it}}\right)^{1-\sigma}} \tag{5}$$

Plugging (5) into (3), we obtain the following gravity equation:

$$M_{ijt} = Y_{it} Y_{jt} \frac{\left(\frac{\tau_{ij} P_{jt}}{P_{it}} + \frac{T_{ij}}{P_{it}}\right)^{1-\sigma}}{\sum_{i=1}^{I} Y_{it} \left(\frac{\tau_{ij} P_{jt}}{P_{it}} + \frac{T_{ij}}{P_{it}}\right)^{1-\sigma}} \tag{6}$$

If we further define the total income of Spain in year t as $Y_{St} = \sum_{i=1}^{I} Y_{it}$ and the share of region i's income in Spain's total income in year t as $S_{it} = \frac{Y_{it}}{Y_{St}}$, equation (6) can be rewritten as:

$$M_{ijt} = \frac{Y_{it}Y_{jt}}{Y_{St}} \frac{\left(\frac{\tau_{ij}P_{jt}}{P_{it}} + \frac{T_{ij}}{P_{it}}\right)^{1-\sigma}}{\sum_{i=1}^{I} S_{it}\left(\frac{\tau_{ij}P_{jt}}{P_{it}} + \frac{T_{ij}}{P_{it}}\right)^{1-\sigma}} \quad (7)$$

where region i's total imports from region j not only depend on the relative incomes of the two regions and on their bilateral trade costs, but also depend on the importing regions' share in Spanish total income and on their average trade costs with respect to all exporting regions.

However, due to the presence of the partner-specific trade cost, the second term of equation (7) contains a non-linearity that calls for a linear approximation so that standard panel data estimation techniques can be applied to the full sample (Baier and Bergstrand, 2009). To approximate that non-linearity, we recall that trade with island regions must necessarily use intermodal (road/sea) or air transport. Both modes of transport carry high specific costs while road-only transport carries high distance-related variable costs. So, over short distances trade costs are higher for intermodal/air transport while over longer distances the cost is higher for road-only transport. Moreover, intermodal/air transport presents higher time costs (Janic, 2007; Wichser et al., 2007). Consequently, intermodal/air transport represents a relatively lower disadvantage in long-distance (international) trade, compared to short-distance (interregional) trade. In the latter case, intermodal/air transport is not competitive with respect to road transport on similar routes, but island regions have no choice (see Hummels (2007) for a discussion of this issue). Therefore, the specific trade cost term represents the additional trade cost incurred by islands due to the use of intermodal/air transport when mainland regions use road-only transport in similar routes. Thus, we take distance non-linearities into account by considering distance ranges that allow a comparison of distance coefficients for mainland and island regions within similar routes.

Furthermore, in line with the basic idea behind gravity models that the intensity with which two partners trade is subject to pull and push factors, we follow Melitz (2007) and assume the total trade cost function TC_{ijt} to be a log-linear function of a set of all observable and unobservable factors that influence trade costs. Accordingly, we also incorporate origin, destination and time fixed effects. Thus we obtain the fully specified trade cost function as follows:

$$TC_{ijt} = \sum_{R=1}^{10} \delta_{1R} d_{ijR} + \delta_2 adj_{ij} + \delta_3 coast_{ij} + \delta_4 islands_{ij} + F_i + F_j + F_t \quad (8)$$

where distance d_{ij} is measured in ten ranges of kilometres covered by road between regional capitals;[10] the dummy adj_{ij} takes value 1 if both regions are adjacent; the dummies $coast_{ij}$

[10] In the case of island regions, sea distance is also measured in kilometres although they are not covered by road. More details on data are provided in section 3.

and $islands_{ij}$ measure the number of coastal regions and the number of island regions in the pair;[11] F_i , F_j, F_t are origin, destination and time fixed effects, respectively.

Log-linearizing equation (7), approximating regional demand by GDP and population, and incorporating the trade cost function (8) to approximate the non-linear term, we obtain the log-linearized reduced-form gravity equation to be estimated:

$$lnM_{ijt} = \alpha + \beta_1 lnY_{it} + \beta_2 lnY_{jt} + \beta_3 lnPop_{it} + \beta_4 lnPop_{jt} + \sum_{R=1}^{10} \delta_{1R} d_{ijR} + \delta_2 adj_{ij} + \delta_3 coast_{ij} + \delta_4 islands_{ij} + F_i + F_j + F_t + \varepsilon_{ijt} \quad (9)$$

where $\frac{1}{Y_{St}}$ is absorbed into the constant term α, common to all years and all country pairs, and into the fixed effect term F_t, and ε_{ijt} is the i.i.d. error term.

2.2. The Blinder-Oaxaca decomposition

Our method to disentangle distance and the "island effect" is based on the Blinder-Oaxaca methodology originally used in labour economics to study the effect of discrimination on wages. The procedure is due to Blinder (1973) and Oaxaca (1973) and it allows decomposing mean differences in any variable based on regression models adopting a counterfactual approach. With this technique, we decompose the bilateral trade differential between two groups of regions into a part that is explained by the regions' characteristics, such as size and distance, and into a residual part that is due to other factors, such as differences in the estimated coefficients associated to the previous characteristics or to unobserved variables. This last term is often used as a measure of discrimination.

More explicitly, we estimate the gravity equation (9) for two groups of regions, group A (mainland regions) and group B (island regions), without imposing the constraint that coefficients are the same for both groups. This is justified by the empirical gravity equation that establishes that trade costs have two components, one that is fixed and another one that is variable. The average trade cost for a firm that wants to export from an island region using a combination of road and sea transport is going to be higher than the average cost that a firm faces when it exports from a mainland region and covers the same distance by road only.

Let's define the expected (average) trade of a region r ($r = \{A, B\}$) as $E[Y_r]$ and the difference of expected trade between the two regions as:

$$\Delta Y = E[Y_A] - E[Y_B] \quad (10)$$

Given that the trade from region r is predicted by a linear gravity model defined as a set of size and friction variables of origin and destination, as represented by equation (9), the method

[11] In this way, these two variables take value 2 if both regions in the pair have coast or are islands, 1 if only one of the trading partners has coast or is an island, and 0 if both trading partners are landlocked or are mainland regions.

gives a precise answer to the question of how much of the difference in expected trade ΔY is explained by differences in size or trade frictions.

The linear gravity model in equation (9) defines a relationship between a trade variable (Y_r) and the regressors (X_r) that can be represented in the following manner:

$$Y_r = X_r' \beta_r + e_r \tag{11}$$

Under the usual assumptions and rearranging, we can decompose the difference of expected trade between the two groups of regions as follows:

$$E[Y_A] - E[Y_B] = [E(X_A - X_B)]'\beta_B + E(X_B)'(\beta_A - \beta_B) + [E(X_A - X_B)]'(\beta_A - \beta_B) = D + C + I \tag{12}$$

Equation (12) shows three main terms. The first term, $D = [E(X_A - X_B)]'\beta_B$, is called endowments and it represents the part of the trade differential that is explained by group differences in the observed characteristics in vectors X_r from the point of view of group B (island regions). In other words, it measures the expected change in Group B's average trade if Group B would have Group A's characteristics. For example, this term will capture the differences in trade due to size and trade frictions that island regions face. The second term, $C = E(X_B)'(\beta_A - \beta_B)$, is called coefficients and it represents the part of the trade differential explained by differences in coefficients. It measures the expected changes in Group B's average trade if Group B had the same coefficients as Group A. For example, this term will capture our hypothesized effect that distance impacts differently on island regions. The third term, $I = [E(X_A - X_B)]'(\beta_A - \beta_B)$, is called interaction and it represents the interaction between characteristics and coefficients. It measures the expected change in the average trade of island regions if these had both the same characteristics and coefficients of mainland regions.

Notice that both the second and third terms from equation (12) include the influence of coefficients on average trade and, as a consequence, both of them are going to be the centre of our interest. Our hypothesis states that differences in trade frictions between Mainland and Island regions (groups A and B, respectively) are going to display a stronger effect in explaining average trade gaps.

The literature has proposed various methods for allocating the interaction term to one of the other two components, so that differences in the variable of interest can be attributed either to differences in characteristics (X's) - called explained difference - or to differences in the coefficients (β's) – called unexplained difference - which is often used as the measure of discrimination. This follows from the two ways of representing the differences in the mean of the variable of interest. On the one hand, the difference between the average trade of Group A and Group B can be expressed by weighting the differences in the X's by the coefficients of Group B, as follows:

$$E[Y_A] - E[Y_B] = [E(X_A - X_B)]'\beta_B + E(X_A)'(\beta_A - \beta_B) \tag{13}$$

On the other hand, the trade variation can also be equivalently expressed with respect to the coefficients of Group A, in which case, the equation becomes:

$$E[Y_A] - E[Y_B] = [E(X_A - X_B)]'\beta_A + E(X_B)'(\beta_A - \beta_B) \qquad (14)$$

In both expressions there are two ways to partition the interaction term, since equations (13) and (14) are actually special cases of the general decomposition defined in equation (12). That is, equation (13) places the interaction into the unexplained part while equation (14) places it into the explained part:

$$E[Y_A] - E[Y_B] = [E(X_A - X_B)]'\beta_B + E(X_A)'(\beta_A - \beta_B) = D + (C + I) \qquad (13')$$
$$E[Y_A] - E[Y_B] = [E(X_A - X_B)]'\beta_A + E(X_B)'(\beta_A - \beta_B) = (D + C) + I \qquad (14')$$

Here it is important to note that the implicit assumption made in equation (13) is that discrimination is directed towards Group A (Mainland regions) and there is no positive discrimination for Group B (Island regions). In equation (14) the assumption is the opposite, that is, there is negative discrimination for Group B (Island regions) but there is no positive discrimination for Group A (Mainland regions). Under our theoretical framework, we are testing whether discrimination runs against Island regions since distance and, likely, other trade frictions are going to have different effects on trade with respect to Mainland regions. In section 4 we present both types of decompositions.

Moreover, Oaxaca and Ransom (1994) suggest the use of a matrix of relative weights, W, as in the following expression:

$$E[Y_A] - E[Y_B] = [E(X_A - X_B)]'[W\beta_A + (I - W)\beta_B] + [(I - W)'E(X_A) + W'E(X_B)]'(\beta_A - \beta_B) \qquad (15)$$

where W and I are the weights of the coefficients of Group A and the identity matrix respectively. When $W = I$, equation (15) becomes equivalent to decomposition (14), and when $W = 0.5$ it is equivalent to the Reimers (1983) decomposition that uses the average coefficients over both groups. However, Neumark (1988) argues that when there is no economic reason to assume that the coefficients of one of the groups should be used as the nondiscriminatory reference model it is preferable to use the coefficients from a pooled regression over both groups. In section 4 we present the results under different assumptions.

3. Features of Spanish regions' trade

In this section Spanish regions' trade is described using two main sources of data for the total international and interregional trade of goods (in millions of euros) of each of the 17 Spanish NUTS2 regions (*Comunidades Autónomas*), in the period 1995-2011.[12] International trade data

[12] Although the data for the autonomous cities of Ceuta and Melilla are available, these cities are excluded from the analysis because their data contains many zeros and they represent less than 3% of Spain's trade flows, so to omit

was obtained from the Datacomex database, which is compiled by Spain's Ministry of Economy and Competitiveness.[13] This database presents trade values and volumes of exports and imports, disaggregated by industry and mode of transport. Interregional trade data was extracted from the C-intereg database, compiled by the Center for Economic Forecasting (*CEPREDE – Centro de Predicción Económica*) at the Autonomous University of Madrid.[14] Interregional trade flows are computed by multiplying the quantities in thousands of tons exported by each region to every other, using each mode of transport, by regional export prices obtained for each year, province (NUTS3 region), mode of transport and product type. The data are cleaned to eliminate the international transit of goods that does not have Spanish regions as origin and/or destination, but it includes goods produced and sold within each region (internal trade).[15] Here bilateral imports are simply the mirror image of bilateral exports. The trade flows data extracted from both international and interregional databases corresponds to the sum of four manufacturing aggregates: agricultural products, intermediate goods, equipment goods and consumption goods. It therefore excludes, for example, energy.[16] However, for simplicity, we shall call "total" to the sum of trade flows for those four manufacturing aggregates.

A potential problem when estimating trade flows between islands and non-island regions is the asymmetric number of zeros, mainly considering the case of non-observable or non-existing flows between the Balearic Islands with landlocked regions (e.g. Madrid). Such limitation is small at the NUTS2 (*Comunidades Autonomas*) level, compared to the NUTS3 (*Provincias*) level data. Moreover, since we are using aggregate trade data, the percentage of zeros in our sample is very low. In particular, for the sample of mainland regions the percentage of zeros is 2.9% (117 out of 4080 observations) while for the sample of island regions the percentage of zeros is 2.2% (12 out of 544 observations).

Data for GDP (at 2008 market prices), population, and GDP's deflator (chained volume index with base 2008) are obtained from the Spanish National Institute of Statistics (*INE – Instituto Nacional de Estadística*). The distance between regional capitals was calculated using the road distance in kilometers, which was extracted from the regional distance matrix provided by the

them spares the problem of the zeros without removing representativeness to the sample. Nevertheless, we initially had carried out regressions with these two cities in the sample and the results are robust to their inclusion (we conclude that it is so precisely because they represent such a low share of flows).

[13] International trade data is used in this section only for comparison and to motivate the analysis. In section 4 our empirical analysis will focus only on interregional trade, that is, trade between Spanish regions.

[14] The C-intereg dataset and its limitations are fully explained in Llano et al (2010). It is important to mention that in the case of the Canary Islands, C-intereg is based on the official data offered by the AEAT (Spain's National Tax Agency), which may include flows between inner regions (non-coastal peninsular regions) and the Canary Islands by ship (implying a transport mode combination); however, this is not possible for the case of the Balearic Islands, where AEAT data is not available. Llano et al. (2011) and Garmendia et al. (2012) also estimate border effects using a similar dataset.

[15] The level of disaggregation of the interregional trade data is restricted due to lack of information or to an insufficient number of observations (e.g., we cannot separate the regional trade flows by mode of transport and origin/destination or before/after crisis years). We can only say that the part of the trade cost that is due to being an island is linked to the need to use air transport or a combination of sea and road transport (intermodal transport).

[16] Energy was excluded to avoid very peculiar situations such as the existence of a refinery in the Canary Islands and the airline hubs (e.g., Air Europa, Air Berlin) located in the Balearic Islands which generate high values of trade flows that distort data.

Spanish National Geographic Institute (*IGN – Instituto Geográfico Nacional*). In the case of the Canary Islands and the Balearic Islands, a combination of road and sea distance is calculated using Google Map's API. This website indicates by default a connection through the closest port. Therefore, we measure the road distance using the minimum distance by road to the closest port with connection to the islands: the ports of Valencia or Barcelona, depending on the destination region, in the case of the Balearic Islands; the port of Huelva in the case of the Canary Islands. In practice this discards that some regions use other ports, but in the absence of data on port use, this is an appropriate sea distance measure since the islands present more intense connections through maritime routes to these ports. We recognize that there should be some measurement error adopting the minimal distance approach but we believe that the error would be compensated with the port choice made by other regions. Moreover, given that we use a range specification for the distance variable, this error is eliminated when two ports are within the same road distance range.[17]

A preliminary look at the trade data already reveals important differences in the trade structure of mainland and island regions (Table 1). On the one hand, island regions sell most of their production internally, that is, within the region (57.9% for island regions against 30.5% for mainland regions). On the other hand, they sell relatively more internationally than to other Spanish regions (the international trade of island regions is 71% of their interregional trade, against 63% for mainland regions). This result demonstrates that island regions are competitive internationally, so that any difficulties they may face in interregional trade should not be attributed to the lack of competitiveness of their products. Moreover, the substantial interregional trade deficit of island regions, driven by their lower share of interregional sales and a higher dependence on interregional imports (52.2% against 37.7% for mainland regions), reveals their trade disadvantage within Spain. This issue assumes greater political economy relevance for these regions, because the small economic size of island regions elevates their trade deficit up to 27.3% of their GDP, whilst they account for only 2.3% and 4.5% of Spain's export and import flows, respectively.

Gravity models suggest that, in the absence of all sorts of barriers typical of international trade such as tariff and non-tariff barriers, or language barriers, the island regions' disadvantage in interregional trade is likely due to the specific costs of trade incurred by these regions. Since distances are greater in international trade, the higher variable costs due to greater distances help to offset the islands' specific costs, which are not present for mainland regions. However, in the shorter distances of interregional trade, variable costs are not high enough to lower the share of specific costs sufficiently and island regions incur a disadvantage that leads to an interregional deficit that is 3.6 times larger than their international trade deficit. The islands' higher share of internal (short distance) trade, their higher interregional (intermediate distance) trade deficit, and their relatively stronger international (longer distance) trade, together seem to suggest the possibility of non-linearities in the trade cost function due to the specific costs incurred in using air transport or an intermodal combination of sea and road transport.

[17] This is the case of Barcelona and Valencia with respect to the Balearic Islands, or Huelva and Cádiz with respect to the Canary Islands.

Table 1. Trade of Spanish regions (1995-2011)

	Total		International			Interregional			Internal		
	Million €	**%Spain**	**Million €**	**%Spain**	**%Geo**	**Million €**	**%Spain**	**%Geo**	**Million €**	**%Spain**	**%Geo**
					EXPORTS						
MAINLAND REGIONS	507,013	97.60%	136,308	98.40%	26.90%	216,170	98.60%	42.60%	154,536	95.50%	30.50%
ISLAND REGIONS	12,059	2.30%	2,110	1.50%	17.50%	2,973	1.40%	24.70%	6,976	4.30%	57.90%
SPAIN	519,359	100.00%	138,456	100%	26.70%	219,160	100%	42.20%	161,743	100%	31.10%
					IMPORTS						
MAINLAND REGIONS	542,494	95.20%	183,370	97%	33.80%	204,588	93.40%	37.70%	154,536	95.50%	28.50%
ISLAND REGIONS	25,860	4.50%	5,387	2.80%	20.80%	13,496	6.20%	52.20%	6,976	4.30%	27%
SPAIN	569,945	100%	189,043	100%	33.20%	219,160	100%	38.50%	161,743	100%	28.40%

TRADE BALANCE

	Total			International		Interregional	
	Million €	**%GDP**	**%Spain**	**Million €**	**%GDP**	**Million €**	**%GDP**
MAINLAND REGIONS	-35481	-4.80%	70.10%	-47063	-6.40%	11582	1.60%
ISLAND REGIONS	-13801	-26.20%	27.30%	-3278	-6.20%	-10523	-20%
SPAIN	-50586	-6.40%	100.00%	-50586	-6.40%	0	0

Sources: DataComex and C-intereg.

Note: The data for Spain is the aggregate of mainland regions, island regions, and the two autonomous cities of Ceuta and Melilla

The empirical evidence provided by the international trade literature suggests that the development of transport by sea has been one of the major integration factors in the world economy (Abe and Wilson, 2009; Behar and Venables, 2010; Hummels, 2007; Limão and Venables, 2001). It is often said that the costs of road transport are much higher than the costs of maritime transport and it can be inferred that the territories with access to the sea have advantages over the rest. However, from our point of view, the island effect has different consequences depending on the distance traveled in the exchanges. When the distance is very long, transport costs can be very low by leveraging economies of scale that come from some technological advances, such as containers or new generations of merchant ships. However, when the distance falls to smaller ranges, the previous technological advances are not as profitable, and exchanges of goods by sea happen through a combination of intermodal transport (roll-roll) combining road transport (usually by truck) with the truckload transportation by ship. Thus, although at long distances (international trade), sea transport is cheaper than land transport, at shorter distances (interregional trade) the variable costs of sea transport may still be lower, but its specific costs become relatively more important.

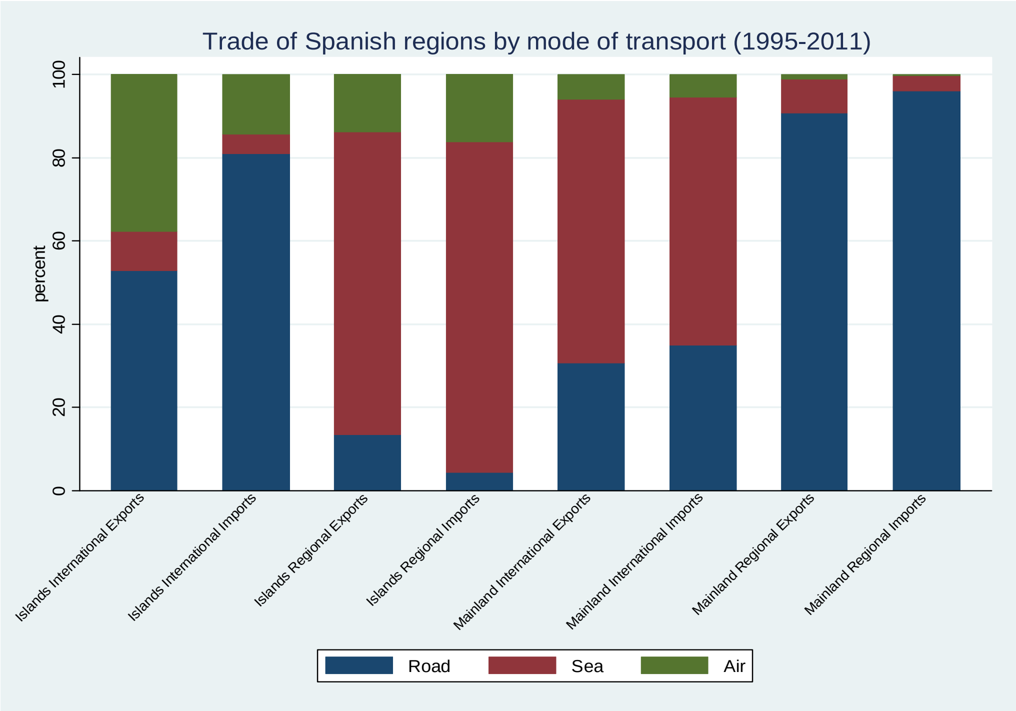
In the absence of comprehensive transport cost data, some indirect evidence regarding the different decomposition of trade costs for mainland and island regions is provided by their relative use of road, sea and air modes of transport in their interregional and international trade. As Figure 2 shows, around 90% of the interregional trade of Mainland regions is carried out by road, even though many of these regions have coast and could trade by sea. This fact evidences that road transport is more suitable for interregional trade, regardless of whether the regions can use sea transport or not. However, the use of sea transport made by island regions in interregional trade is roughly comparable to that of mainland regions in international trade (60-70%). This is because mainland regions are able to use only road transport in their trade with other mainland regions, whilst island regions have to use an intermodal combination of road and sea transport, which bears high specific costs in terms of fees and taxes, besides the time lost in transport.[18]

Moreover, although for both region groups the use of air transport is higher in international than in interregional trade, island regions' use of this transport mode is relatively higher in both cases.[19] This is thus indirect evidence that higher specific trade costs are incurred by island regions in shorter distance (lower variable cost) trade compared to longer distance (higher variable cost) trade, with the consequence that the average trade cost is relatively higher for island regions, especially at intermediate distances such as those of interregional trade. Combining this information with the importance of internal trade, it appears justified that the impact of distance may be non-linear and depends on the distance range of trade. This non-linearity will be formally tested in the following section.

[18] In the islands' trade, the higher share of road (sea) in international (interregional) trade is two sides of the same coin: the transport from Palma to Paris (Valencia) has a high share of road (sea) and a low share of sea (road).

[19] Although in the case of Spain the higher use of air transport in interregional trade made by island regions is determined by tourism-related activities, it can also be seen as the lack of alternative of using road-only transport.

Figure 2



Sources: DataComex (international exports and imports) and C-intereg (interregional exports and imports).

In conclusion, it can be said that, in the case of Spain, island regions trade relatively more internally and internationally than with other regions in the same country. They also show a strong dependency from regional imports and this, together with their higher focus on international export markets, causes a high and persistent interregional trade deficit. At the same time, the performance of island regions in international markets reveals that they do not lack competitiveness, but instead their interregional deficit is due to higher trade costs springing from, on one hand, the impossibility of using solely road transport and, on the other hand, a greater average distance from markets and suppliers.

4. Empirical results

4.1. Gravity estimation results

The first step of our empirical strategy consists of the estimation of equation (9) for interregional exports and imports. Moreover, the gravity models for each type of trade flow are estimated for two different samples: Group A (Mainland) when the exporter/importer is a mainland region and the trade partner is either a mainland or an island region; and Group B (Islands) when the exporter/importer is an island region and, again, the trade partner is either a mainland or an island region. The dependent variable is either the total value of exports or the total value of imports from region j to region i in year t, expressed in logarithms. The explanatory variables included in the baseline model are exporter and importer GDP and population (also expressed in logs), a set of pairwise dummy variables for adjacency, coast and island, and distance, as defined in section 2.1. The island variable measures the number of islands in the pair, therefore taking values from 0 to 2. The non-linearity of distance is captured by splitting it in 10 ranges of equal frequency (deciles) denoted as $DistR2_{ij}$-$DistR10_{ij}$.[20] This increases the flexibility of our approach and allows us to capture a richer pattern of transport costs across several distance ranges.

The estimation results are fairly standard within the gravity model literature (see Table 2). In general, the model explains 81%-85% of the variation in interregional trade for island and mainland regions. As predicted by the gravity equation, the economic size of the regions, measured in terms of GDP and population, matters to explain both exports and imports, although differences exist depending on the flow and the sample considered. In the mainland region sample, the GDP of the reporter (mainland) regions is significant while it is not significant for the partner (mainland and island) regions. In the islands sample, the GDP of the partner (mainland and island) regions affects both exports and imports positively, while that of the reporters (islands) is not relevant. On the other hand, more population in the exporter region reduces both exports and imports for mainland regions, while more population in the exporter region increases exports from island regions but reduces its imports. Therefore, economic size has a differentiated impact on mainland and island regions. In particular, since the island regions are relatively small, their GDP is not relevant to explain their interregional trade flows but the economic size of the partner region matters.

As expected, trade costs affect bilateral trade negatively, but with non-linearities. In all the ranges considered, distance coefficients are negative and highly significant. Moreover, the coefficients of distance ranges for Mainland's exports and imports are very similar, meaning that transportation costs have symmetric effects on the trade direction. However, for the Islands' exports and imports there is more asymmetry, since Islands' exports are generally more affected by distance than imports. Regarding the shape of the distance effect, we observe for most columns that the absolute magnitude of the coefficients increases with distance ranges in ranges $DistR2_{ij}$-$DistR9_{ij}$, but at a decreasing rate.[21] In the case of Island's imports they increase up to the distance of 664 km and show more variability for higher distances. Therefore, distance affects trade following a U-shaped pattern.[22] Importantly, when comparing the distance ranges coefficients between mainland and island regions we see that they are larger within the same distance range for island regions. The distance gap tends to be higher at the lower ranges and lower at the higher ranges for both exports and imports. This is clearly supporting our view that the impact of trade costs is higher for island regions, especially at shorter distances, and that the impact is non-linear.

[20] Distance ranges are defined as ten dummies that comprise the following distance intervals: $R1_{ij}$=[0, 174), $R2_{ij}$=[174, 277), $R3_{ij}$=[277, 390), $R4_{ij}$=[390, 458), $R5_{ij}$=[458, 549), $R6_{ij}$=[549, 664), $R7_{ij}$=[664, 739), $R8_{ij}$=[739, 936), $R9_{ij}$=[936, 1608) and $R10_{ij}$=[1608, 2404]. For example, the first dummy $DistR1_{ij}$ takes the value 1 when the distance between region *i* and region *j* ($R1_{ij}$) is lower than 174 km and takes value 0 otherwise. The omitted category in the regression analysis is the dummy for the shortest distance range $DistR1_{ij}$.

[21] Here is important to note that $\mathrm{DistR10}_{ij}$ refers to trade with the Canary Islands.

[22] A quadratic approximation to distance non-linearity is described in subsection 5.3.

Table 2. Gravity model estimates

	Group A: Mainland regions		Group B: Island regions	
	(1)	(2)	(3)	(4)
	Exports	Imports	Exports	Imports
GDP_{it}	1.371***	1.630***	-3.310	6.424
	(0.363)	(0.574)	(8.955)	(5.357)
GDP_{jt}	0.746	1.106	5.698**	3.836**
	(0.557)	(0.685)	(2.267)	(1.590)
Pop_{it}	-1.480***	-0.567	14.06*	2.283
	(0.336)	(0.430)	(7.467)	(4.412)
Pop_{jt}	0.00962	-1.033*	-3.302	-2.730**
	(0.468)	(0.587)	(2.011)	(1.304)
$DistR2_{ij}$	-1.212***	-1.177***	-3.368***	-2.555***
	(0.174)	(0.187)	(0.798)	(0.335)
$DistR3_{ij}$	-1.646***	-1.675***	-4.182***	-2.356***
	(0.164)	(0.173)	(0.798)	(0.335)
$DistR4_{ij}$	-1.953***	-1.948***	-4.790***	-4.312***
	(0.169)	(0.183)	(0.798)	(0.335)
$DistR5_{ij}$	-2.305***	-2.303***		
	(0.182)	(0.189)		
$DistR6_{ij}$	-2.857***	-2.696***	-5.206***	-5.680***
	(0.207)	(0.221)	(0.829)	(0.380)
$DistR7_{ij}$	-2.691***	-2.642***	-5.376**	-4.239***
	(0.215)	(0.228)	(2.270)	(0.347)
$DistR8_{ij}$	-3.180***	-3.149***	-5.411***	-5.114***
	(0.211)	(0.223)	(0.888)	(0.776)
$DistR9_{ij}$	-3.220***	-3.325***	-5.846***	-4.303***
	(0.240)	(0.249)	(0.824)	(0.386)
$DistR10_{ij}$	-3.109***	-4.832***	-3.879***	-3.879***
	(0.296)	(0.323)	(0.522)	(0.219)
$Adjacency_{ij}$	0.134	0.183		
	(0.110)	(0.118)		
$Coast_{ij}$	0.654	2.325***	3.854***	1.786
	(0.823)	(0.818)	(1.260)	(3.541)
$Island_{ij}$	-0.165	-2.009***	-0.890	-1.610
	(0.537)	(0.545)	(3.830)	(2.214)
Constant	8.912*	0.783	-167.9**	-82.42
	(5.188)	(6.386)	(71.35)	(59.69)
Observations	4,281	4,251	503	533
R-squared	0.841	0.860	0.808	0.847

Notes: Robust standard errors clustered at region-pairs in parentheses. Significance levels are denoted by *** p<0.01, ** p<0.05 and * p<0.1. Dependent variables, GDP and Population variables are in logs. Fixed effects (i, j, t) are included in all regressions (but omitted for brevity). "DistR" denotes distance ranges.

The adjacency dummy would be likely to pick up the effect of short distance on trade. It is well known in the literature that geographic aggregation produces strong economic effects (i.e. artificially large border effects and underestimation of distance on trade). Llano et al. (2011), using NUTS3-level data (a more disaggregated geographical unit than the one used here), show that trade among mainland regions is highly geographically concentrated in Spain. As shown by Hillberry and Hummels (2008) and Llano et al. (2011), the high concentration of trading also explains the high positive value of the adjacency coefficient. If trade data has a low geographical disaggregation, the adjacency variable might pick the strong non-linear relationship between trade and distance. However, the fact that we are already accounting for distance non-linearities in the distance variable itself renders the adjacency variable irrelevant. Island regions do not have adjacent regions, and the use of road-only transport is not possible. Consequently, as estimated, the distance effect becomes more relevant for these regions. For mainland regions, the adjacency dummy is not significant since the short distance effect is presumably reflected in short distance ranges. Alternatively, this could be because we use data at the NUTS2 level or because we include islands in our sample.

The island dummy is significant for imports and presents the expected negative sign in both samples. Since trade costs are mainly faced by the importing region, mainland regions import less from island regions, but it is not relevant whether the exports are destined to another mainland region or to an island region. Besides, as stated before, island regions are highly dependent on imports from the mainland. In our data, the results just mean that Balearics and Canary Islands trade less with each other than with mainland Spain, which is to be expected within the gravity framework. At the same time, the coast dummy is significant (and positive as would be expected) only in exports from island regions.[23] In this way, our coast variable controls for the islands' need to use a sea port in their interregional exports.

To sum up, Table 2 results show that the relationship between distance and trade is different for island and for mainland regions, as shown by the higher magnitude of the distance coefficients for the islands, as well as by their higher exports to coastal regions. We believe that these differences in the coefficients are due to the higher trade costs faced by the islands, as well as to their specific character.

4.2. Results from the Blinder-Oaxaca decomposition

Once the gravity model is estimated for the two groups, the second step consists of applying the Blinder-Oaxaca decomposition to the gravity results of Table 2. The results for the general decomposition, applied separately to exports and imports, are presented in Table 3. The results show that Island regions trade less with other Spanish regions than Mainland regions. In the case of exports, the trade difference is 3.1 log points while in the case of imports the difference is lower but still significant, meaning that the lower bound for the regions trade difference is 1 log point. The decomposition also shows that the difference due to endowments is 7.5 for exports and 5.7 for imports, the difference due to coefficients is 5.3 and

[23] Note that *Coast* and I*sland* variables are defined as the number of coastal and island regions in the pair, respectively. So, in the mainland sample coast takes the values 0, 1 or 2 while island takes the values 0 or 1.

-2.4 for exports and imports respectively, and the differences due to the interaction are -8.7 for exports and -2.3 for imports.

These results reveal that differences in endowments and in interactions are the most important explanations for the trade gap between regions but they are not significant. Nevertheless, the differences in coefficients -that contain the effect of interest- are significant. However, due to the fact that the effects of interactions partially cancel out the other effects (especially for exports) as they are of opposite signs, the interaction effect could be potentially relevant in the analysis of discrimination.

Table 3. General decomposition

	Exports	Imports
Mean prediction for		
(A) Mainland Regions	5.907***	5.694***
	(0.100)	(0.120)
(B) Island Regions	2.820***	4.692***
	(0.408)	(0.329)
Difference (A)-(B) due to	3.087***	1.002***
	(0.420)	(0.350)
Endowments	7.490	5.667
	(5.490)	(3.684)
Coefficients	4.269***	-2.388**
	(0.838)	(0.933)
Interaction	-8.673	-2.277
	(5.533)	(3.777)

Notes: Robust standard errors clustered at region-pairs in parentheses. Significance levels are denoted by *** p<0.01, ** p<0.05 and * p<0.1. Dependent variables are in logs.

To make the interpretation clear, the analysis on Table 3 shows that if Island regions had the same characteristics as Mainland regions, the mean exports at Islands would increase by 7.5 log points while imports by 5.7 log points. That is, the differences in endowments are more important to explain differences on expected exports than differences on imports. Similarly, if Island regions had the same coefficients as Mainland regions, exports would rise at Islands by 4.3 log points and imports would be reduced by 2.4 log points. The interaction between endowments and coefficients has the same sign as the coefficient effect and opposite to that of the endowment effect, which makes unclear what the joint effect of differences in coefficients would be on the trade outcome, since the interaction component embodies a part of the explanatory power of coefficients.

Before entering the discrimination analysis, we show in Table 4 the separate effect of detailed grouped variables such as Size, Distance and Other Frictions. Size indicates the joint prediction made by GDP and population from i and j, whilst Distance represents the join effect of distance from i to j. Other Frictions contains all (i, j) terms - the dummy for adjacency, the coast variable, and the islands variable – as well as the fixed effects (i, j, t). Focusing only in the coefficients, the results show that the Distance effect is positive for both exports and imports but is statistically significant only for exports while the effect of Other Frictions is mostly negative for exports and positive for imports. Furthermore, although the coefficient of Distance is non-significant for imports, the interaction is significant and part of this effect is likely a pure coefficient effect.

Table 4. Detailed decomposition

	Endowments	**Coefficients**	**Interaction**
Panel A. Exports			
Size	2.761	-176.0**	-2.928
	(1.865)	(70.19)	(1.867)
Distance	0.387	1.391**	0.216
	(0.398)	(0.578)	(0.363)
Other Frictions	-0.599	-4.833	0.542
	(3.632)	(4.233)	(3.638)
Fixed effects	4.941**	6.880	-6.503***
	(2.088)	(5.477)	(2.162)
Total	7.490	4.269***	-8.673
	(5.490)	(0.838)	(5.533)
Panel B. Imports			
Size	2.190	-92.40	-1.966
	(1.610)	(59.60)	(1.584)
Distance	0.677***	0.221	0.732**
	(0.263)	(0.396)	(0.299)
Other Frictions	0.994	0.489	0.253
	(0.958)	(3.625)	(0.992)
Fixed effects	1.806	6.098	-1.296
	(1.722)	(4.121)	(1.839)
Total	5.667	-2.388**	-2.277
	(3.684)	(0.933)	(3.777)

Notes: Robust standard errors clustered at region-pairs in parentheses. Significance levels are denoted by *** p<0.01, ** p<0.05 and * p<0.1. Dependent variables are in logs. The Coefficient column sum up to the Total effect when we introduce the differences accounted by the intercept.

From the previous decomposition, we are unable to disentangle what part of the interaction term is attributed to the Endowment difference and what part to the Coefficient difference. Therefore, in Table 5 we perform a discrimination analysis as described in Section 2.2. In Column (1) we present the results assuming that discrimination runs only against Group B (Island regions), as in equation (14). In Column (2) we assume that discrimination is directed towards Group A (Mainland regions). In Column (3) we employ Reimers (1983) weight, which is a simple average of coefficients of both types of regions. This is a priori our preferred assumption because it doesn't imply a strong commitment regarding what region is most trade discriminated. Last, in Column (4) we present Neumark's (1988) approach consisting of the use of the coefficients from the pooled model.

The results show that the unexplained component is positive and significant in three out of four columns. In terms of our model, this implies that exports at Island regions should increase if coefficients where the same as in Mainland regions. Under our preferred assumption (i.e. third Column), exports should rise by 1.5 log points at islands. Regarding imports, the unexplained component is positive and also significant at three out of four columns, meaning that imports would rise at Islands if coefficients were the same as at Mainland regions. Again, under our preferred assumption, imports should increase by 0.6 log points at islands if both regions would face the same transport costs. The asymmetric effect of distance on exports and imports is difficult to interpret but it is possibly related to the reaction of exporter firms in the islands to find more advantaged locations.

There are other noticeable unexplained effects captured by other variables. For instance, size discrimination operates with a negative sign, meaning that Islands would trade less if size coefficients would be the same as in Mainland regions. However, this effect is only significant for exports. Apart from distance, Other Frictions seem to exert an important effect for exports in two out of four columns, while the fixed effects seem to have an important influence in the case of –two out of four columns- imports.

In sum, the results in Table 5 show for exports and imports respectively that in 3 out of 4 decomposition methods the unexplained component of distance is significant. This suggests that there is a wide support for our hypothesis. Moreover, in the case of exports, the unexplained distance component is almost significant at 10 per cent level in the fourth method, which reinforces our earlier conclusion.

Table 5. Detailed Oaxaca decomposition

	(1)		(2)		(3)		(4)	
	Explained	**Unexplained**	**Explained**	**Unexplained**	**Explained**	**Unexplained**	**Explained**	**Unexplained**
Panel A. Exports								
Size	-0.167	-176.0**	2.761	-178.9**	1.297	-177.5**	-0.133	-176.0**
	(0.170)	(70.19)	(1.865)	(71.61)	(0.940)	(70.90)	(0.219)	(70.18)
Distance	0.604***	1.391**	0.387	1.608**	0.495*	1.500***	1.044***	0.950
	(0.213)	(0.578)	(0.398)	(0.631)	(0.263)	(0.577)	(0.256)	(0.586)
Other Frictions	-0.0566	-4.833	-0.599	-4.290**	-0.328	-4.561*	0.568	-5.457
	(0.288)	(4.233)	(3.632)	(1.915)	(1.824)	(2.735)	(1.020)	(3.783)
Fixed effects	-1.562***	6.880	4.941**	0.376	1.689	3.628	0.158	5.159
	(0.566)	(5.477)	(2.088)	(4.316)	(1.082)	(4.811)	(0.432)	(5.590)
Total	-1.182	4.269***	7.490	-4.403	3.154	-0.0671	1.637	1.450
	(0.822)	(0.838)	(5.490)	(5.471)	(2.785)	(2.769)	(1.201)	(1.194)
Panel B. Imports								
Size	0.224	-92.40	2.190	-94.36	1.207	-93.38	0.217	-92.39
	(0.144)	(59.60)	(1.610)	(60.99)	(0.824)	(60.29)	(0.149)	(59.64)
Distance	1.409***	0.221	0.677***	0.952***	1.043***	0.587*	1.023***	0.607*
	(0.297)	(0.396)	(0.263)	(0.314)	(0.237)	(0.324)	(0.243)	(0.342)
Other Frictions	1.247***	0.489	0.994	0.743	1.121**	0.616	4.039***	-2.302
	(0.370)	(3.625)	(0.958)	(4.514)	(0.530)	(4.064)	(0.731)	(3.766)
Fixed effects	0.509	6.098	1.806	4.802*	1.158	5.450	-0.944	7.552**
	(0.672)	(4.121)	(1.722)	(2.769)	(0.930)	(3.388)	(0.865)	(3.845)
Total	3.390***	-2.388**	5.667	-4.665	4.528**	-3.527*	4.336***	-3.334**
	(0.923)	(0.933)	(3.684)	(3.671)	(1.909)	(1.899)	(1.347)	(1.335)

Notes: Robust standard errors clustered at region-pairs in parentheses. Significance levels are denoted by *** p<0.01, ** p<0.05 and * p<0.1. Dependent variables are in logs. Unexplained columns sum up to the Total effect when we introduce the differences accounted by the intercept.

5. Robustness checks

5.1. The role of tourism

The islands' tourism-dependency may cause distortions in the estimations. Naturally being islands makes the Balearic and Canary islands more attractive for sand and sun tourism, thus increasing internal demand during the high season and increasing imports during those months. Distance and tourism-dependency are likely to be positively correlated. In our analysis we are already controlling for supply-side factors of comparative advantage using fixed effects, and the potential demand by two time-varying size measures: GDP, that includes the domestic revenue produced by all the sectors in the economy, including the tourism sector; and permanent population, that controls for the number of regular consumers, part of which are workers in the tourism sector, but excludes the tourist visitors. It can be argued that our measure of population does not capture the sudden peak that potential visitors represent for demand during the tourism season and, therefore, we should try to control for this effect.

In practical terms we should add the "appropriate" measure of floating population to the importer regions but not to the exporter ones. Introducing such a measure will reduce the explanatory power of the distance variable *only* if both variables, floating population and distance, are not orthogonal.[24] On the other hand, this raises some measurement issues to tackle. First, we have to consider that tourist arrivals for each region should be weighted according to the average length of the stay. Second, we have some data availability restrictions because data on international tourist arrivals by destination region are available only from 2001. For domestic tourism, data are available since 2006. Third, there would be an underreporting of the domestic tourism flows since it is much harder to detect people's movement when there are no border crossings. Moreover, in the longer term, tourist visits may also increase exports to their regions of origin. Fixed effects would pick this up, but it can also be isolated in the way described above in the exports equation.

Therefore, we have extended our baseline regressions for imports and exports to include the floating population of the reporter region ($FloatPop_{it}$) in logs as previously described.[25] We present the results in Table A.1 in the appendix. As we mentioned, there is a reduction in the number of observations and, importantly, the coefficients are not significant in either equation. We also have repeated the decomposition analysis in Table A.5. The Oaxaca decomposition now shows a higher relevance of the unexplained component of the distance variables in comparison with the baseline decomposition. Alternatively, in Table A.2 in the appendix we have also estimated the baseline gravity model with the share of tourism in GDP ($TouShare_{it}$) and in Table A.5 the corresponding decomposition analysis.[26] The results in the

[24] This is unlikely given that all Spanish regions experience a huge increase in the number of visitors during the summer season.

[25] We use data of international tourist arrivals and average length of the stay obtained from the inbound tourism survey (FRONTUR) elaborated by the Spanish Institute of Tourism Studies. Floating population is calculated as (inbound tourists*length of the stay)/365.

[26] The share of the tourism sector (hospitality) on GDP is obtained from the Spanish Regional Accounts elaborated by the Spanish National Institute of Statistics (INE).

gravity equation are quantitatively the same as this variable is not significant either and the unexplained components of the Oaxaca decomposition remain unaltered. This means that our results are robust to the introduction of tourism and the structure of the economy does not affect our baseline results.[27]

5.2. Distance non-linearities

As mentioned in section 4, the empirical literature shows that most economic interactions take place in proximity. This is a plausible explanation of why distance coefficients are found to increase over distance ranges. This formulation is also justified by the data, as shown in section 2, where it can be observed that Spanish island regions trade more at shorter and longer distances, but less at intermediate distances. Moreover, previous studies, such as Hillberry and Hummels (2008), use the distance and the square of the distance to address the non-linear relationship between trade and distance. These authors show that most shipments occur at very short distances. Our baseline approach is aimed at capturing this feature of the data in a more flexible way, but an alternative is to impose a specific functional form to distance non-linearities using a quadratic distance term.

When distance enters the regressions as a second order polynomial, the distance linear term is highly significant with a negative sign, and the quadratic term is positive and significant, as shown in Table A.3 in the appendix. The adjacency variable remains not significant. Therefore, distance affects trade following a U-shaped pattern, as had already been found using distance ranges: the negative distance coefficients increase (in absolute value) with distance ranges, but at a decreasing rate. These results show that distance matters less in short and long distances (internal and international trade), but it matters the most in intermediate distances (interregional trade).

Since our interest is in comparing the coefficients for two samples, we show that the distance effect is much stronger at island regions, but only for exports. The unexplained components of the distance variables displayed in Table A.5 are not significant. This is due to the fact that the quadratic specification only approximates the actual shape and that the range-by-range does a better job in estimating the non-linearities found in the data. Importantly, the range approach makes it very clear that the distance coefficients for island regions are larger than the coefficients of mainland regions.

5.3. Multilateral resistance

Anderson and van Wincoop (2003) generalized the basic gravity model to incorporate multilateral resistance terms. Their incorporation is important because they account for the attractiveness of trading with third regions. Holding bilateral trade costs constant, less difficulty in trading with third regions implies a higher ratio of bilateral to multilateral trade

[27] We also have performed an additional check (not reported) introducing origin-time and destination-time dummies in the gravity model and obtained higher distance coefficients for island regions.

costs and thus a decrease of bilateral trade. So bilateral trade could be affected negatively (positively) when trade with third parties is facilitated (hindered). In their work, multilateral resistance was approached by price indices for the exporter and the importer regions. Although their approach was difficult to implement empirically, the two important points they made were that the multilateral resistance measures should be weighted averages of characteristics of all trading partners in the sample and should be time-varying.

We check the robustness of our baseline results using an approach to multilateral resistance that is derived directly from equation (7), whose denominator term models multilateral trade resistance through all variables that also influence the bilateral resistance to trade, weighted by regional income shares. Since these shares are time-varying, we can overcome the bias present in estimations that solely rely on country (pair) fixed effects to proxy for the multilateral resistance terms (Spies and Marques, 2009). In order to be able to apply standard panel data estimation techniques, the multilateral resistance variables are written as a Taylor series approximation as proposed by Baier and Bergstrand (2009):

$$MR_{ijt} = \frac{1}{I}\sum_{i=1}^{I} Z_{ijt} + \frac{1}{J}\sum_{j=1}^{J} Z_{ijt} - \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Z_{ijt} \qquad (16)$$

where Z_{ijt} is the product of regional income shares and each regional characteristic (distance, adjacency, coast and island). The first two terms represent the multilateral trade resistances of the respective trading partners. Holding bilateral trade costs constant, a rise in these terms implies a lower ratio of bilateral to multilateral trade costs and thus a boost of bilateral trade. The last term, however, resembles the country resistance to trade and as such, lowers the trade value between every pair of regions. To give an example, for the distance variable this means that a higher distance of the trading partners *i* and *j* towards all other regions in the sample increases region *i*'s imports from *j*, whilst a high country distance (everyone is far away from everyone) lowers trade between every pair of regions. In the case of dummy variables, the relevant economic interpretation is given by the proportion of 1 versus the proportion of 0 values of the dummies for the various pairs of regions. The opposite interpretation of the multilateral and world resistance terms holds, of course, true for trade-stimulating factors, like adjacency.

Accordingly, we have introduced five multilateral resistance terms in the baseline augmented gravity model (Table A.4).[28] The baseline results remain robust to the incorporation of the multilateral resistance variables defined above: bilateral distance remains significant in all ranges and it still has higher effects for islands. The multilateral resistance terms themselves are not always significant and the excessive magnitude of some coefficients makes us suspect that the introduction of these variables creates a colinearity problem. When decomposing the effect of distance (Table A.5) we also find that the unexplained component is very high and significant. That is, our results are robust to the introduction of multilateral resistance terms.

[28] To account for the non-linear effect of distance we have split the multilateral resistance term of distance in two components, a linear one and a quadratic one.

6. Conclusions

The main objective of this paper is to disentangle the effect of distance and of sea borders on interregional trade involving island regions. To that end, we consider the case of Spain, a country with two island regions: Balearic Islands and Canary Islands. The impact of land borders, as political borders that negatively affect the trade of adjacent regions located in different countries, has been extensively studied in the international trade literature. On the contrary, the role of sea borders, as geographical borders that impact negatively on the trade of non-adjacent regions located in the same country, remains unstudied. Sea borders produce similar negative effects upon trade as land borders, as they originate specific costs, and thus require a modification of the trade cost function to reflect that. However, land political borders imply trade policy and language barriers, whilst sea geographical borders raise the need to pay for additional logistic costs, as well as the time-loss inefficiency related to the use of an intermodal (truck plus ship) transport mode. Whereas in the former case barriers can be reduced through free trade agreements, this instrument is not available in the latter case.

The empirical strategy used in the paper consists of two different stages. Firstly, a gravity model for interregional exports and imports is estimated for two different groups: Island regions and Mainland regions. Then, a Blinder-Oaxaca decomposition is applied to the gravity estimation results in order to disentangle the distance and border effects for those regions, net of all other factors controlled for in the gravity estimations.

We present evidence of the relevance of the Island effect as a special border effect, since island regions are at a substantial disadvantage compared to mainland regions. We disentangle the channels through which the Island effect determines trade flows among regions and evaluate their relative importance in explaining trade gaps with respect to Mainland regions. In particular we extend a gravity model to include different trade costs for Islands and estimate the separate effect of distance and other trade frictions for Island and Mainland regions.

Our findings suggest that Island regions are unevenly affected by regional characteristics but more importantly by the estimated coefficients associated to distance. This is consistent with our hypothesis that there are specific trade costs that Island territories are subject to. Moreover, our results suggest that, among the different variables that reduce trade, distance is by far the most important variable explaining the trade gap among different types of regions. It presents a non-linear (U-shaped) behaviour, with more negative coefficients for Islands at each distance range, which validates the presence of an island-specific cost of trade, especially at intermediate distances.

The results leave no doubt that island regions are at a substantial disadvantage in trade compared to mainland regions. Islands naturally reacted to their higher specific cost by substituting intermediate distance (interregional) trade with short distance (internal) and long distance (international) trade. Since the geographical particularity of island regions, which prevent the use of road-only transport, cannot be avoided, there is room for policy intervention by subsidizing the transport of merchandise in and out of islands and/or by improving the sea and air transport infrastructure. In Spain, for example, the domestic transport of residents is already publicly subsidized in both island regions. In particular, residents in the islands receive a 50% discount in their domestic air and boat tickets. Similar measures could be taken for merchandise transport.

Finally, although the paper uses data for Spain, it establishes a methodology that is transferable to other trade contexts where there are both islands and continental regions. Such extensions, in the European and in the world context, are left for future research.

References

Abe, K. and J. Wilson (2009), 'Weathering the Storm: Investing in Port Infrastructure to Lower Trade Costs in East Asia', World Bank Policy Research Working Paper 4911.

Anderson, J.E. (2011), 'The gravity model', Annual Review of Economics, 3, 133-160.

Anderson, J.E. and van Wincoop, E. (2003), 'Gravity with Gravitas: a Solution to the Border Puzzle', American Economic Review, 93 (1), 170–192.

Armstrong, S. (2007). 'Measuring Trade and Trade Potential. A Survey', Asia Pacific Economic Papers No. 368, Australia-Japan Research Centre.

Baier, S. and Bergstrand, J. (2009), 'Bonus Vetus OLS: A Simple Method for Approximating International Trade-Cost Effects Using the Gravity Equation', Journal of International Economics, 77(1), 77-85.

Behar, A. and Venables, A. (2010), 'Transport Costs and International Trade', Handbook of Transport Economics. Ed. A. de Palma, R. Lindsey, E. Quinet and R. Vickerman. Edward Elgar, 2010.

Blinder, A. S. (1973), 'Wage Discrimination: Reduced Form and Structural Estimates', The Journal of Human Resources, 8, 436–455.

Clarke, X., Dollar, D. and Micco, A. (2004), 'Port efficiency, Maritime transport cost and bilateral trade', Journal of Development Economics, 75, 417-450.

Chen, N. (2004), 'Intra-National versus International Trade in the European Union: Why do National Borders Matter?', Journal of International Economics, 63, 93-118

Eilat, Y. and Einav, L. (2004), 'The determinants of international tourism: a three dimensional panel data analysis', Applied Economics 36(12), 1315-1328.

Evans, C. L. (2003), 'The Economic Significance of National Border Effects', The American Economic Review, 93 (4), 1291-312.

Feenstra, R.C. and J. Romalis (2014), 'International Prices and Endogenous Quality', Quarterly Journal of Economics, 129, 477-527.

Fourie, J. and Santana, M. (2013), 'Cultural Affinity and Ethnic Reunion'. Tourism Management, 36, 411-420

Garmendia, A., Llano, C. and Requena-Silvente, F. (2012), 'Networks and the Disappearance of the Intranational Home Bias', Economic Letters, 116, 178-182.

Gil-Pareja, S., Llorca-Vivero, R., Martínez Serrano, J. A. and Oliver-Alonso, J. (2005), 'The Border Effect in Spain', The World Economy, 28(11), 1617–1631.

Gil-Pareja, S., Llorca R. and Martínez J.A. (2006), 'The Impact of Embassies and Consulates on Tourism', Tourism Management 28, 355-360.

Gil-Pareja, S., Llorca R. and Martínez J.A. (2007), ‘The Effect of EMU on Tourism’, Review of International Economics 15, 302-312.

Head, K. and Mayer, T. (2000), ‘Non-Europe: The Magnitude and Causes of Market Fragmentation in Europe’, Review of World Economics, 136 (2), 285-314

Head, K. and Ries, J. (2008), ‘FDI as an Outcome of the Market for Corporate Control: Theory and Evidence’, Journal of International Economics 74(1), 2-20.

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

Hummels, David (2007), ‘Transportation Costs and International Trade in the Second Era of Globalization’, Journal of Economic Perspectives, Vol. 21(3), 131–154.

Janic, M. (2007), ‘Modelling the Full Costs of an Intermodal and Road Freight Transport Network’, Transportation Research Part D, 12, 33–44

Limão, N and A Venables (2001), ‘Infrastructure, Geographical Disadvantage, Transport Costs and Trade’, World Bank Economic Review, 15 (3), 451-479.

Llano C.; Minondo A. and Requena-Silvente F. (2011), ‘Is the Border Effect an Artefact of Geographical Aggregation?’, The World Economy, 34, 10, 1771-1787.

Llano. C., Esteban. A., Pulido. A. and Pérez. J (2010), ‘Opening the Interregional Trade Black Box: The C-intereg Database for the Spanish Economy (1995-2005)’, International Regional Science Review, 33, 302-337.

Marques, H. (2010), ‘Migration Creation and Diversion in the EU: Any Crowding-Out Effects from the CEECs?’, Journal of Common Market Studies, 48(2), 265-90

Marques, H. (2011), ‘Asymmetries in Heterogeneous Integrated Areas: Evidence from Sectoral Trade between Old and New EU Members’, Journal of International Trade and Economic Development, 20(1), 5-29

Marques, H. and Metcalf, H. (2005), ‘What Determines Sectoral Trade in the Enlarged EU?’, Review of Development Economics, 9(2), 197-231

Marques, H. and Metcalf, H. (2006), ‘Ending Restrictions to Migration from the New EU Member Countries: Sectoral Trade and Real Wage Effects’, Contemporary Economic Policy, 24(2), 287-99

Márquez-Ramos, L., Martínez-Zazoso, I., Pérez-García, E. and Wilmsmeier, G. (2007), ‘Determinantes de los Costes de Transporte Marítimos. El Caso de las Exportaciones Españolas’, Información Comercial Española, 834, 79-93.

McCallum, J. (1995), ‘National Borders matter: Canada-US regional trade patterns’, American Economic Review, 85(3), 615-23

Melitz, J. (2007), ‘North, South and Distance in the Gravity Model’, European Economic Review, 51, 971-91.

Neumark, D. (1988), ‘Employers’ Discriminatory Behavior and the Estimation of Wage Discrimination’, The Journal of Human Resources, 23, 279–295.

Novy, D. (2013), ‘Gravity Redux: Measuring International Trade Costs with Panel Data’, Economic Inquiry, 51(1), 101-21.

Oaxaca, R. (1973), ‘Male-Female Wage Differentials in Urban Labor Markets’, International Economic Review, 14, 693–709.

Oaxaca, R. L., and M. R. Ransom (1994), ‘On Discrimination and the Decomposition of Wage Differentials’, Journal of Econometrics, 61, 5–21.

Papazoglou, C., Pentecost, E. and Marques, H. (2006), 'A Gravity Model Forecast of the Potential Trade Effects of EU Enlargement: Lessons from 2004 and Path-dependency in Integration', The World Economy, 29(8), 1077-1089

Reimers, C. W. (1983), 'Labor Market Discrimination Against Hispanic and Black Men', The Review of Economics and Statistics, 65, 570–579.

Rosselló, J. and Santana, M. (2014), 'Recent Trends in International Tourist Climate Preferences: a Revised Picture for Climatic Change Scenarios', Climatic Change, 124. 119-132

Sanso, M., Cuairán, R. and Sanz, F. (1993), 'Bilateral Trade Flows, the Gravity Equation, and Functional Form', The Review of Economics and Statistics, 75(2), 266-275

Sanz, F. (2000), 'A Kalman Filter-Gravity Equation Approach to Assess the Trade Impact of Economic Integration. The Case of Spain (1986-1992)', Weltwirtschaftliches Archiv, 136(1), 84-110.

Santana, M., Ledesma, F.J., Pérez, J.V. and Cortés, I. (2010), 'Does a Common Currency Promote Countries' Growth via Trade and Tourism?', The World Economy, 33(12), 1811-1835

Spies, J. and Marques, H. (2009), 'Trade Effects of the Europe Agreements: a Theory-based Gravity Approach', Journal of International Trade and Economic Development, 18, 11-35

Wichser, J. Weidmann, U. and Nash, A. (2007), 'Strategies for Increasing Intermodal Freight Transport between Eastern and Western Europe', Association for European Transport and contributors.