Response to Reader 4 (Carlos Llano)'s Comments

Thank you very much for your careful review of our paper. We appreciate your constructive and thoughtful feedback. We specially thanks all the issues related to data availability and limitations. Adding discussion on the dataset will make the paper clearer for the readers.

Below we describe in detail our responses to your comments. We are happy to revise the paper along the lines you propose and as described in the responses below.

Overview

0. *This paper studies the border effect in Spain between 1995-2011 trying to disentangle the role of distance and the fixed costs induced by being an Island region. The paper combines international and interregional aggregate flows for the 17 Spanish regions NUTS 2 (autonomous communities). 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 more to the lack of adjacency imposed by the sea border rather than to the higher average distance.*

Authors' response: It is important to clarify that international trade data are used only to motivate the paper. However, the model estimated in the empirical section only considers interregional trade between Spanish regions. This point is clearly stated in footnote 13.

General comments

The paper deals with an interesting topic, discussing to what extent Island regions face advantages/disadvantages when enrolling in international and interregional trade, with regards to non-island regions of the same country. The results are reasonable and suggest and interesting line for further developments. The use of the Blinder–Oaxaca decomposition, borrowed from the labor economics literature can be also considered as a methodological contribution. Along to these virtues, the paper also presents some drawbacks. Some of them can be easily applied in a revised version of this paper. Others, in my opinion, will require some improvements in the data used and maybe can be considered in further investigations. Next you will find some comments and suggestions.

Specific comments

Aim and motivation:

In general the aim and motivation is ok. However it could be improved saying something about the following points:

1. The paper suggests the potential disadvantage of Island regions with regards to "interregional trade", taking the mainland regions as reference. Although this is reasonable, it will be interesting to try to connect that with the expected results found in border effect and international trade literature, as well as on the complex relation between trade and transport logistics. For example: 1) in international trade literature, "land-locked" regions are usually considered the most disadvantaged locations. Ships are [still] the main mode for long distance deliveries, and ports have also enjoyed mayor geographical advantages throughout history as strategic enclaves for trade, inducing economic agglomeration. Is it possible that the "island effect" considered here (as a disadvantage), which plays in the opposite direction than having access to the sea, may just take place for small or remote Islands? Is this effect the same for the Canary Island and the Balearics? Having this in mind, apart from decomposing the effects with regards to Non-Island Spanish regions, maybe the reference group should be the "landlocked" regions in Spain. 2) It is important to consider that Catalonia and Comunidad Valenciana (coastal regions) have a kind of "geographical monopoly" over the Balearic Islands. Something similar happens with Andalusia with respect to the Canaries. Thus, if firms have decided to locate in these coastal regions with the aim of serving the Island on distance, probably these coastal regions should not be considered in the reference group, since their exporting capacity (interregionally) is inflated "over the average".

Authors' response: We also perceive that the relationship between trade and transport costs is a complex one and it should be clearly discussed at the motivation. Therefore the points made below are now discussed in section 3.

Regarding point 1), the empirical evidence provided by international trade suggests that the development of transport by sea has been one of the major integration factors in the world economy. It is often said that the costs of road transport are much higher than the costs of transport by sea. 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 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. Furthermore, as presented in Figure 2, interregional trade for Mainland regions (many of them with coast) is carried out mainly by road (around 90%) while island regions mainly use maritime transport. This fact evidences that road transport is more suitable for interregional trade, regardless of whether they are coastal regions or not. This restricts the use of landlocked regions as the reference group to better capture the differential dependence on maritime trade.

On the other hand, the reader is right that splitting the results for the Canary Islands and the Balearic Islands is of interest. However, the lack of variability for distance ranges when dealing separately with both archipelagos impedes the identification of several distance parameters with precision. This is why we prefer to consider both island regions simultaneously in the analysis and include in the gravity equation the variable "island" that measures the number of islands in the pair.

Regarding point 2), it is natural that some regions have specific advantages to trade with the islands and that this feature generates agglomeration in particular coastal cities. It

happens similarly to Madrid, a landlocked region with a strong commercial relationship with the Islands. Whatever the agglomeration sources are, we take them as given and stress that our interest relies on distinguishing trade cost parameters between island and non-island regions to capture the effect of intermodal transport (truck+ship), which is the major transport mode with the islands, while it is absent in trade exchanges between mainland regions. In any case, our "coast" variable controls for being landlocked, so the "island" effect is already net of trading with a landlocked partner.

2. *It would be also nice to include some comments about the complex relation between "transport mode-trade-trade costs" for Islands, Coastal and Inner regions with respect to trade. For example: 1) Although it is true that for the case of the two Spanish Island regions considered here, the ship is the main transport mode, aircraft also plays a role, which is more relevant for longer distances and more expensive products. For this flows using aircraft, the right distance is the straight line, deliveries are quicker but the transport cost per Ton*km increases; 2) Moreover, although the combination of two transport modes rises the transportation cost (road-ship), a large part of interregional deliveries with the two Islands use the "roll-and-roll" strategy, where the truck is loaded into the ship, reducing the logistic cost in the two maritime terminals (Puertos del Estado provides some information about this); 3) Canary Island enjoy a specific tax systems that is supposed to promote trade (interregional as well as international), something that can compensate higher transport costs…*

Authors' response: Regarding point 1) and point 2), we have extended the discussion on the trade cost by mode of transport and added some data regarding their relevance. Although we cannot explicitly state the percentage of merchandise that goes through each mode of transport or through each port, by using distance ranges we are at least able to know which geographical ranges are there in the data for each region.

Regarding point 3), we completely agree that Canary Islands already receives compensation for being an island region while the Balearic Islands do not and this situation is not being controlled for in our analysis. That is, part of the effect of distance should be stronger if the Canary Islands' specific tax system did not exist. The counterfactual where the Canary Islands trade in similar conditions to the Balearics is not feasible. So, our analysis could be interpreted as a lower bound of the effect of distance on trade for the islands.

3. *A robust check using 2 time periods (before/after the crisis) may be interesting.*

Authors' response: Year fixed effects are included in all regressions. So, time-specific events, such as the current crisis, are being controlled. Moreover, we have run regressions with region-specific shocks (results not reported for brevity) using region-time fixed effects and the results show that trade costs coefficients are still larger for island regions.

Flow Data:

4. *The paper combines international (Datacomex) and interregional (C-intereg) flows for a long period. It seems that this paper just uses the aggregate flows. I think that this is right as a first step, but the paper should be clearer about this, since there are some implications for such option. For example, two points are worth mentioning: 1) C-intereg's aggregate flows usually include "R-16. Production and Distribution of Water, Gas and Electricity", a sector that is not included in the international equivalent flows, and which may introduce some bias for the Island regions; 2) In the case of Canary Islands, C-intereg is based on the official data offered by the AEAT, 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 Balearic Islands, where AEAT data is not available. I think that such limitations can just be solved using and improved version of the dataset, which is out of author`s hands at this moment. However, having this in mind, I will recommend: 1*) including a robust analysis where the Balearic and the Canary Islands are analyzed separately. Not just the different quality of the data, but the geographical and economic differences between both also support this option; 2*) including a reference to the paper where the C-intereg dataset is explained (Llano et al, 2010) which explains in detail such limitations. Also including a reference to other paper estimating border effects and using a similar dataset (Llano et al, 2011; Garmendia et al, 2012) will raise the comparability of the results.*

Authors' response: We agree that more discussion on data availability and limitations is necessary.

Regarding recommendation 1) it is important to clarify that our data excludes sector R-16 to avoid such problems and that several checks have been made to ensure that other omitted geographical and economic differences are not driving the results. We have made this point clearer in section 3 and added section 5 with robustness checks.

Regarding recommendation 2) we especially thank the reader for the references provided that are now referenced in footnote 14. Finally, we now explicitly explain in section 3 that the trade flows used (both international and interregional) correspond to the sum of four manufacturing aggregates (agricultural products, intermediate goods, equipment goods and consumption goods), as we published on July 18.

Regarding recommendation 1*), see explanation given on comment 1 from **Aim and Motivation**.

Regarding recommendation 2*), these references have been added to the revised version of the paper.

5. *A potential problem when estimating 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 inner regions (i.e. Madrid). Such limitation is small at the Nuts 2 level, a point that can be used by the authors with regards to some suggestions raised about the use of Nuts 3 level data. Anyway, commenting this potential limitation, and providing the % of zero flows for the Islands/Non-Islands when describing the dataset may elude further criticism. Robust checks using PPML will also be interesting.*

Authors' response: As mentioned by the reader, since we are using aggregate trade data the % of zeros in our sample is very low. In particular, for the sample of mainland regions the % of zeros is 2.9% (117 out of 4080) while for the sample of island regions the % of zeros is 2.2% (12 out of 544). This information was added to section 3.

6. *Figure 2 provides statistics about transport mode in regional trade. The labels and source are not clear. You should emphasize more clearly if this graph just considered interregional or international flows, since the mode-mix is clearly different. What about "train".*

Authors' response: The clarity of this figure was improved following the reader's recommendations. Although there is data for transport by train, it represents a very low share of interregional trade and so we excluded it, as well as postal packs, etc.

Regressors:

7. *In addition to the regressors included in the paper [standard] it would be interesting to include the followings:*
   a. *Split the Island dummy for considering the case of being an exporter or an importer Island region/country. It is interesting to see if the transport mode fixed cost is working similarly in the two directions.*
   b. *Split the Island dummy for the two Island regions. As commented, in addition to the singularities of the data, interregional exports from the Canary Island are mainly driven by "the refinery located in Tenerife" (as well as agricultural products. i.e: banana, tomatoes…) which implies a complex logistical links with other enclaves, not always associated with pure distance, GDPs or Population (i.e. intra-firm trade between the refineries located in different regions in Spain: Cartagena, Tarragona, La Coruña, Bilbao…).*
   c. *As another reviewer has suggested, it would be nice to know additional details about the distance variable used in the paper. What about interacting the distance with the "Island dummy" in order to capture the "obliged" transport mode combination that implies trading with an Island?*

Authors' response: Again, following the reader's recommendations some robustness checks were carried out.

Regarding point a), we carry out the analysis separately for exports and for imports. Since these two flows are the mirror image, the islands dummy is showing the effect of being an exporter or an importer in each equation.

Regarding point b), although splitting dummy variables for the two island regions is an interesting robustness check that we cannot incorporate for the reasons stated above, it is important to highlight that our exports and imports variable excludes energy. That is, it is generated by adding up exports/imports on agriculture, intermediate inputs, equipment and consumption goods as we now state in section 3. Energy was excluded precisely to avoid the problem with the refinery in the Canary Islands and the airline hub (AirEuropa, Air Berlin,…) located in the Balearic Islands. Nevertheless, it is true that the differences observed could be a consequence of the different specialization of both island regions that comes up when data are more disaggregated.

Regarding point c), It is important to clarify that most trade with Islands is carried out using maritime transport (as can be seen from Figure 2) and that the Blinder-Oaxaca decomposition picks up this difference without relying on the restrictiveness of the dummy approach (by interacting the island dummy with distance). Intermodal transport has high specific costs while road-only has high iceberg costs. So, over short distances trade costs are higher for intermodal transport while over longer distances the cost is higher for road-only transport. Moreover, intermodal transport presents higher time costs (Janic, 2007; Wichser et al, 2007). Consequently, intermodal transport represents a relatively lower disadvantage in long-distance trade (i.e international), compared to interregional trade. Therefore, intermodal transport is not competitive with road transport on similar routes, but island regions have no choice given the low substitutability of both transport modes. Hummels (2007) and related literature discuss this issue.

Nevertheless, the specific trade cost term represents the additional trade cost incurred by islands due to the use of intermodal or air transport. This cost is due to the existence of a sea border for islands. We have added a detailed discussion on this to the revised version of the paper to improve its motivation.

References:

Llano. C., Esteban. A., Pulido. A., 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.

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

Garmendia, A., Llano, C. and Requena-Silventre, F. (2012) ""Network and the disappearance of the intranational home bias,"" Economic Letters 116, 178-182.