Referee 2 report on
"Islands in trade: disentangling distance from border effects"
submitted to Economics: The Open-Access, the Open-Assessment E-Journal

Thank you very much for your careful review of our paper. We appreciate your constructive and thoughtful feedback. 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.

General comment:

The paper studies the effect of sea border on trade between regions in Spain. In order to do that the authors estimate a trade equation for Spanish regions including two island regions and apply a Blinder-Oaxaca decomposition to disentangle the distance and the border effect. The paper answers an interesting question what is the impact of a sea border on inter-regional trade. In general the idea is interesting. There are however several issues that I would like the authors to address.

Response to referee's comments

1. *My first set of remarks is very general. The two island regions are tourism economies. It has several implications for the paper. First of all, their main focus are tourism activities and therefore they do not produce many export products. As a consequence, they highly rely on imports of consumer goods from Spain for the tourism sector. Second, if tourists are included the population of these regions is much larger (for example according to http://www.spanish-fiestas.com/ more than 4 mln people visited Balearic islands in 2013). It explains the high trade deficits of this regions. It also may be a possible explanation for the insignificant distance effect for island's region exports. I would guess that a lot of ships may leave islands regions with less products than they brought there. Hence, the authors should try to control for tourism sector for example by taking into account the number of tourists or their spending in the regression analysis.*

Authors' response: The referee is absolutely right. We now discuss this issue in the newly created robustness checks section 5 (subsection 5.1).

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. We are already approaching the potential demand by two time-varying size measures: GDP, that includes the domestic revenue produced by all 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.[1] 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 problems 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. However, there is a reduction in the number of observations and the coefficients are not significant in either equation. Moreover the Oaxaca decomposition shows now a higher relevance of the unexplained component of the distance variables. Alternatively, we have also estimated the baseline model with the share of tourism in GDP and the results are quantitatively the same as this variable is not significant either. Moreover, we 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. This means that our results remain robust to the introduction of tourism and the structure of the economy does not affect our results.

2. *My second set of remarks concerns the **descriptive statistics section**. I find this part very interesting; however there are several issues that are worth to be explained.*

• *Table 1 is really difficult to read and therefore I am not sure what the authors mean when saying for example that island regions sell more domestically (within the region or within Spain?). I also did not quite understand the following sentence: "they sell relatively more internationally than to other Spanish regions (the interregional trade of island regions is 71% of their international trade, against 63% for mainland regions)", as the second part of it does not really prove the first part.*

Authors' response: These sentences have been corrected. Internal trade refers to regional production domestically consumed. This issue is clarified in the text.

• *There is an important issue that we can read from Figure 2. Island international imports and exports rely mostly on road transport while their interregional counterparts on sea transport. This indicates that there might be a problem with the data. Islands international trade should solely rely on sea transport. However, I can imagine that most of trade flows operate via mainland regions and then they shipped to island regions (and recorded as being transported using roads). In this case the interregional and international statistics are recorded differently and they should not be directly compared. I would like the authors to elaborate more on that issue.*

Authors' response: The referee is correct. The higher shares of road (sea) in international (interregional) trade are 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). Please note that Figure 2 shows percentages. This is now explained in footnotes 18 and 19.

---

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

• In addition, the authors say that island regions have a disadvantage in interregional trade since they have to use both road and sea transport. This is not necessarily true since sea transport is found to be much cheaper than land transport. For example Limao & Venables (2001) find that an extra 1000km distance raises costs by seven times more if the distance is overland than if it is maritime. [See more references on the transportation costs in the references. The authors should include more references from this literature.]

References:
Abe, K & J Wilson (2009), 'Weathering the Storm: Investing in Port Infrastructure to Lower Trade Costs in East Asia', World Bank Policy Research Working Paper 4911.
Behar, A & A Venables (2010), 'Transport Costs and International Trade', Department of Economics Discussion Paper Series (Ref: 488)
Hummels, D (2007), 'Transportation Costs and International Trade in the second era of Globalization', Journal of Economic Perspectives 21 (3), 131–154.
Limão, N & A Venables (2001), 'Infrastructure, Geographical Disadvantage, Transport Costs and Trade', World Bank Economic Review, 15 (3), 451-479. Table").

Authors' response: The referee is correct, there is a literature showing that ground transport is at a disadvantage with respect to maritime transport, which we now cite. For example, Limão and Venables (2001) show that ground transport costs are seven times higher than maritime transport costs per distance unit. This estimate is difficult to generalize, since they only focus on the costs of shipping a standard container from Baltimore to other distant ports. However, the containerization of international shipments requires investing in expensive port infrastructure and both developments are only profitable when the volume of operations from port to port is high enough. Our hypothesis deviates from this standard point of view, because the sea distance covered by interregional shipments is lower and the scale of the market is not large enough. We now discuss this crucial issue along the paper.

Our model allows us to capture these features. For example, iceberg costs are likely to be proportional to distance. At long distances (international trade), sea transport is cheaper than land transport. However, at shorter distances (interregional trade), although the iceberg costs of sea transport may still be lower, its island-specific costs become relatively more important due to the islands' need to combine road and maritime transport.

3. *My third set of remarks concerns the regression analysis.*

• Including the quadratic distance effect creates the following issues. First of all, using logarithms controls already for non-linearity of the distance. Second the authors do not necessarily prove that distance matters the most at intermediate distances. For the island regions they find that distance is not significant (see my first set of remarks for a possible explanation) for exports, and the coefficients point to an U-shaped curve for imports. However, the size of the coefficients indicates a minimum at about 900 km. This distance in case of the islands regions (especially Canary Islands) is not an intermediate distance. Second as far as mainland regions are concerned, the inverse U-shaped curve reaches the maximum at respectively at 30 km for exports and 50 km for

imports. Since there are not such short distances between regions it does not show the non-linearity. Hence, using traditional log of distance seems to be more appropriate.

Authors' response: We now motivate in greater detail the introduction of non-linear distance effects, which we now introduce in section 4 using distance ranges, and confirm their existence in section 5 with a quadratic polynomial (subsection 5.2).

The empirical literature shows that most economic interactions take place in proximity. This is a plausible explanation of why distance coefficients are found to change with distance ranges (e.g. the magnitude of distance coefficients increases with the distance to trading partners). Our approach is aimed at capturing this feature of the data in a flexible way using distance ranges and a robustness check using a quadratic distance term. Moreover, the distance variable and its quadratic term are now considered without using logs for the reasons explained next.

Our current results show that the relationship between distance and trade is different for island and for mainland regions, and that this fact can be explained by the existence of different trade costs, but the results could be contaminated by the way distance is treated. We have checked whether our results are sensitive to this by inserting distance into the gravity equation in two different ways. First, we have used our distance variable without logs entering as a second order polynomial, as in the baseline model of the Discussion Paper version. In both imports and exports equations now the distance linear term is highly significant with a negative sign, and the quadratic term is positive and significant. Interestingly, now the adjacency variable turns out insignificant, as presumed. So, we conclude that it is better not to log. Second, we have split the distance variable in ten equally frequent intervals (i.e. decile ranges) instead of treating it as a continuous variable. This increases the flexibility of our approach and allows us to capture a richer pattern of transport costs across several distance ranges. In all the ranges, distance coefficients are negative and highly significant and the adjacency variable is no longer significant. These results show more clearly that for each distance range coefficients are larger for island regions. The Oaxaca decomposition shows again that most of the trade gap among regions is due to the additional transportation costs that island regions suffer.

Although in both checks we find that distance affects trade following a U-shaped pattern, by estimating distance coefficients we find 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. Therefore, the latter estimation is now our baseline model and quadratic distance is discussed as a robustness check in subsection 5.2. of the revised version.

Importantly, our main message remains valid after this change in the specification: the distance coefficients for island regions are larger than the coefficients of mainland regions. Hence, an important part of the trade gap is still explained by differences in coefficients and not by physical distance.

4. **Minor comments**

• *Empirical strategy part is too long (especially the part on the gravity equation) and would be more logical if it was placed just before the empirical results part.*

Authors' response: The empirical strategy part is now shorter as some text was moved to the robustness checks section 5. The main purpose of presenting the gravity model is to justify the existence of non-linearities, but we agree that it is a fairly standard model.

    • *Abbreviations in Table 1 are not clear.*

Authors' response: The table's note has been extended to better explain the table's contents.

    • *If the data for the regions of Ceuta and Melilla is available, why not including them in the estimations (especially given the fact that they have a sea border with other Spanish regions)?*

Authors' response: The reason is that there are many zeros for these two cities and they represent less than 3% of Spain's trade flows, so to omit 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 suppose that it is so precisely because they represent such a low share of flows). This is now explained in footnote 12.

    • *Typos in the x axis of the Figure 2.*

Authors' response: This figure has been amended.

    • *It would be more logical to change the order of columns in table B1 since the authors refer to the first two ones as group B and to the last two ones as group A.*

Authors' response: Table B1 is now Table 2 as it has been placed in the main text in response to another set of comments. The modification suggested by the referee has now been incorporated too.

    • *Table 5 is in the middle of the conclusion.*

Authors' response: This is no longer the case.