Response to Reader 3 (Asier Minondo)'s Comments

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.

Specific comments:

1. *The authors introduce the fixed effects in transport using a quadratic distance. In particular, they use the log of distance and the square of the log of distance. I am not sure whether the specification should use logs. Previous studies, such as Hillberry and Hummels (2008), use the distance and the square of the distance and not the log of distance and the square of the log of distance to address the non-linear relationship between trade and distance. These authors show that most shipments occur at very short distances. Llano et al. (2011), using province-level data (a more disaggregated geographical unit than the one - autonomous communities-used by Groizard et al.), show that trade among mainland regions is highly geographically concentrated in Spain. This pattern is different to that found by Groizard et al. for island regions. 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.*

   Authors' response: We now motivate in greater detail the introduction of quadratic distance. 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 a quadratic distance term. Moreover, taking the reader's point, distance and its quadratic term are now considered without using logs and the references suggested have been incorporated.

   The reader's concern regarding the variable "Adjacency" is well taken because this variable is 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). For instance, Llano et al. (2011) point out that "If the sub-national unit is geographically large, trade between sub-national units may not pick the sharp reduction in value that happens at short distances." 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, while looking for changes in the adjacency coefficient. 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.

2. *The authors use the Blinder-Oaxaca decomposition to measure how trade would change if islands did not bear the costs of changing the transport mode and were located closer to the main centers of economic activity in Spain. However, authors should bear in mind that, as pointed out by Anderson and van Wincoop (2003), in the counterfactual situation (islands become mainland regions, or distance is reduced) the multilateral resistances would change as well. These changes should be taken into account when comparing the actual and the counterfactual situation, and I am not sure whether the Blinder-Oaxaca procedure controls for this. Anderson and van Wincoop (2003) use Gauss to estimate the actual and counterfactual multilateral resistances. As the coding in Gauss is not straightforward, Baier and Bergstrand (2009) develop an alternative, and easier to implement, procedure to estimate the multilateral resistances in the counterfactual situation.*

Authors' response: We agree that we need to check whether our results are sensitive to multilateral resistance (MR) and we now do that in subsection 5.3. Although some MR terms are significant, the baseline estimate results remain valid. Besides, we suspect that the introduction of MR terms creates colinearity problems.

The point we make in the paper is that Islands trade less than mainland regions for two reasons: they have different characteristics and they face different "prices" associated to those characteristics (i.e. a great penalty to transport merchandises with origin or destination in the islands). To isolate those effects we generate 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 transport 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). When we perform the second counterfactual we impose that each region has its own coefficients, including its own multilateral resistance terms, because any change in those terms would transfer some explanatory power to the characteristics, and thus, the decomposition will be misleading. On the other hand, trying to simulate what the multilateral resistance terms of regions would be if they would have the same characteristics, such as distance, adjacency or size is technically challenging and intuitively misleading for our purpose (i.e. the MR of the Balearic Islands or the Canary Islands would depend on the exact location we place the islands on the mainland territory).

3. *The authors estimate the regression with origin-region fixed-effects, destination-region fixed effects and time fixed-effects. I would like to see whether the estimations are altered if higher dimensional fixed effects are introduced. As the authors claim, the correct estimation of the regression demands origin+time and destination+time fixed effects. This would preclude the estimation of some variables, such as GDP or population; but, it would not preclude the estimation of the key variables for the analysis, such as distance and island.*

   Authors' response: We have checked whether the estimations are altered when introducing origin-time and destination-time fixed effects as an alternative way to control for origin and destination price changes over time and because these terms are interpreted by some authors as multilateral resistance terms. Effectively some variables cannot be estimated, but the key variables such as distance and island are estimated. After fitting the model under several distance specifications the coefficients for distance are still larger for island regions. These results, not shown for space reasons, reinforce our main message.

4. *Authors use a combination of road and sea distance to calculate distance between the mainland regions and island regions. The authors do not explain whether they choose the distance from the origin region to the region which is closest to the island region for the road interval and then add the sea distance. I wonder whether Spanish mainland regions always use the same ports to trade with the Balearic islands and the Canary Islands. For example, it might be the case that most firms in Galicia and most firms in Madrid use the port of Valencia to transport their goods to the Canary Islands. If that were the case, distances had to be adjusted.*

   Authors' response: 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. This is the case of Barcelona and Valencia with respect to the Balearic Islands, or Huelva and Cádiz with respect to the Canary Islands (see footnotes 17 and 18). Nevertheless, we now explain more clearly our methodology to compute sea distance and state the approximate distance range of trade by Balearic and Canary islands.

5. *Garmendia et al. (2012) show that social networks and business networks also explain why some Spanish regions trade more with other Spanish regions.*

Authors' response: The point we make in the paper is that distance restricts to a greater extent trade between regions than cannot substitute maritime transport by other more competitive transport modes for similar distances. As long as other sources are orthogonal to distance our study remains valid. The trade networks literature does indeed suggest that the strength of trade networks is negatively correlated with distance but positively correlated with trade flows. A possible way to control for the presence of networks would be to use a time-varying variable such as the number of residents originating in the partner region present in the reporter region. Unfortunately such variable would be correlated with the error term and an IV estimation would have to be carried out. However, the trade networks literature suggests that the strength of trade networks is negatively correlated with distance but positively correlated with trade flows. This would reinforce our results as the consideration of networks would widen the trade gap between Island and Mainland regions and the varying strength of networks is itself a result of distance. This potential channel although interesting is beyond the scope of our research.

References

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