RESPONSE TO THE REFEREE

First of all, we would like to thank the referee for his/her attention, comments and suggestions.

Referee's report: **The Effect of Technological Innovation on International Trade. A Nonlinear Approach** - by Laura Márquez-Ramos and Inmaculada Martínez-Zarzoso

The paper explores an interesting topic of international trade, i.e., the relationship between the technological achievement of a country and the value of its export. The originality of the approach steams from the adopted data sources – namely the TAI index developed by UNDP – and the robustness of the results is carefully assessed through a number of different econometric specifications.
I found the paper very interesting and generally well written, however I think a number of points need to be addressed before it will be publishable.

1) The emphasis on the non linearity of the relationship is probably excessive, given that it is assessed only by the mean of a quadratic term, which is quite a common practice in literature.

Answer:
We follow there the usual practice in the literature.

2) The sample selection is carefully explained but not motivated – why only 13 countries? Furthermore, the claim that the selected countries are "representative" may appear arbitrary, and a wider sample may significantly increase the robustness of the results, in the light of the following point.

Answer:
When all the countries are included as exporters, the sample size increases considerably when using sectoral trade flows. This is the main reason why we chose a smaller sample of countries that are representative of the different stages of development and the different specialisation patterns. Following the referee's recommendation, we also tested the relationship with a sample of 65 countries using aggregated trade flows. See the next answer.

3) The analysis is carried on at country **and** sector level but the reason for that is not clear to me – why not running the analysis also on the pooled sample at the country level (assuming the number of countries is increased)? To the extent that almost all the explanatory variables – and all the relevant ones – are grouped at country level, this seems a sensible option. The additional information on differentiated trade patterns according to the Rauch taxonomy can be exploited by aggregating export flows at that level, as done in the robustness tests, or by introducing interaction terms (e.g. sector variables interacted with TAI)

Answer:
As an additional robustness check, we did the same analysis for aggregated trade flows using a larger sample of countries in the revised version of the paper. We used a 65-country sample which uncovers the countries selected in Figure A.1 in the paper with a

*maximum of 4160 (65*64) bilateral trade flows by using data for 1999. The coefficients obtained for technological innovation are higher than those obtained when using disaggregated data. However, they were also higher when the same analysis was performed for the aggregated trade flows of the 13-exporter sample. As in the disaggregated analysis, a "U-shaped" relationship was found between exports and the diffusion of old innovations, whereas an inverted "U-shaped" relationship was seen between exports and the diffusion of recent innovations, and between exports and human skills. Otherwise, the opposite results were found for the dimension of creation of technology since an inverted "U-shaped" relationship was noted between exports and creation of technology with the aggregate samples.*

Therefore two columns have been added to Table 2 (Table 2. The effect of technological innovation on international trade).

	IV_65 countries	***IV_13 countries aggregated***
Exporter's TAI	*13.38****	*10.95****
	(9.03)	*(3.10)*
Exporter's TAI (square)	*-7.43****	*-9.38****
	(-4.86)	*(-2.71)*
Importer's TAI	*10.26****	*1.51***
	(7.52)	*(2.37)*
Importer's TAI (square)	*-7.33****	-
	(-5.10)	
Exporter's creation of technology	*9.24****	*2.69***
	(13.88)	*(2.26)*
Exporter's creation of technology (square)	*-11.43****	*-3.32**
	(-8.55)	*(-1.80)*
Importer's creation of technology	*5.72****	*3.31****
	(7.76)	*(3.06)*
Importer's creation of technology (square)	*-8.06****	*-4.04***
	(-5.36)	*(-2.00)*
Exporter's diffusion of old innovations	*-7.73****	*1.81****
	(-3.39)	*(3.38)*
Exporter's diffusion of old innovations (square)	*9.18****	-
	(6.28)	
Importer's diffusion of old innovations	*-2.36*	*-1.26**
	(-1.36)	*(-1.84)*
Importer's diffusion of old innovations (square)	*3.94****	*2.43****
	(3.45)	*(3.61)*
Exporter's diffusion of recent innovations	*7.29****	*1***
	(11.46)	*(1.99)*
Exporter's diffusion of recent innovations (square)	*-4.09***	-
	(-5.27)	
Importer's diffusion of recent innovations	*3.81****	*1.38****
	(6.06)	*(4.02)*
Importer's diffusion of recent innovations (square)	*-2.09****	-
	(-2.61)	
Exporter's human skills	*11.19****	*6.49****
	(10.94)	*(3.50)*
Exporter's human skills (square)	*-6.21****	*-5.17****
	(-6.98)	*(-3.21)*
Importer's human skills	*6.89****	*3.54****

	(7.05)	*(3.25)*
Importer's human skills (square)	*-4.22****	*-1.96**
	(-4.76)	*(-1.93)*
1-U Theil	*0.85*	*0.91*
R-squared	*0.76*	*0.86*
Number of observations	*1,895*	*799*

4) The fact that the four components of TAI and the synthetic index (i.e. the average of the four components) are included in the same specifications is probably causing important collinearity problems.

Answer:
All the columns in Table 2 of the paper show the results obtained when estimating the gravity model which was augmented with each technological variable: creation of technology, diffusion of recent innovations, diffusion of old innovations, human skills, and the TAI. As the referee points out, to include all the technological variables in the same regression can lead to important collinearity problems. Therefore, five regressions were estimated with each methodology (OLS, PPML, Harvey and IV). Table 2 shows the coefficients of the technological innovation obtained in each estimated regression in order to compare the results across specifications and methodologies.
In the paper we point out: "Then, TAIi and TAIj are the technological variables measuring technological innovation in the exporting and importing countries. To analyse the individual effect of the different dimensions composing the TAI on international trade, four additional regressions were derived from Equation (3) where TAI can be substituted by each of its four dimensions. In order to analyse the existence of a non-linear relationship between technological innovation and international trade, two additional terms are included in the model, $(TAI_i)^2$ *and* $(TAI_j)^2$*. Then, this index is decomposed into its four dimensions and the model is again estimated with the two additional terms in each dimension." (page 12)*

5) A number of issues arise from the IV specification. First, there is no way to see if they are strong, as first stage results are not mentioned. Second, I understand that two variables are used as instruments, while the endogenous variable are five (the TAI index and its four components), which is clearly leading to an underidentification problem. Third, the exogeneity of the instruments is dubious, as they can be easily correlated with many omitted variables which may affect the export flows. In order to test exogeneity, I suggest to resort to the Hausman test (for overidentified specifications), rather than the one used in the paper, which does not seem to be very popular. Fourth, once the IV specification is correctly specified, it may be interesting to compare the results with the OLS ones, but this requires the specification to be identical, which is not the case in the paper.
6) Generally, omitted variables bias and reverse causality is a major issue, especially given the cross-section nature of the data, thus more work on points 5 may be highly rewarding in terms of robustness of results.

Answer: These analyses are performed in Section 5.3.
Although gravity models are, in most cases, estimated using OLS, this specification does not account for the existence of causality between technological innovation and trade. Since this is a potential problem that will lead to a misspecification of the estimated

model, the possible endogeneity of technological innovation in the gravity equation is analysed in this paper (see Section 5.3). In this case, technological innovation will be correlated with the error term, while the OLS estimates will be biased and inconsistent.

In order to test the presence of endogeneity, a Hausman test is performed in the 65-country sample. The purpose of this test is to indicate whether there is correlation between technological innovation and the error term in the augmented gravity model. The null hypothesis is that there is no correlation and, therefore, that OLS provides consistent and efficient estimates. If this is true, the IV estimates should be similar to the OLS estimates.

To estimate by IV, the use of a set of instrumental variables that are correlated with technological innovation in countries, but not with the error term of Equation (3)[1]*, would be desirable. Average research and development expenditure (% of GDP) and average public spending on education (% of GDP) in the period 1994–1998 were selected as technological innovation instruments. The selection of the instrumental variables is based on Eaton and Kortum (2002). These authors suggest that a country's level of technology is related to its stock of past research effort, and that a higher stock of human capital allows a country to absorb more ideas from abroad.*

The version of the Hausman test proposed by Davidson and MacKinnon (1993) is also applied to validate the results. In a first step, technological innovation (TAI_i) is regressed on all the exogenous variables and on an instrument to obtain the residuals. Then in a second step, the augmented gravity model is estimated including the residuals from the first regressions as an additional variable. When R&D expenditure is used as a proxy for technological innovation, the residuals of those regressions are not significant in the augmented gravity model. Since the coefficients on the first-stage residuals do not significantly differ from zero for R&D expenditure, the test indicates that there is no endogeneity problem. Therefore, the OLS estimation is consistent. However, when public spending on education is included as a technological innovation instrument, the residuals of those regressions are significant in the augmented gravity model. Thus when public spending on education is included as an instrument, the test accepts the hypothesis of endogeneity.

Finally, whether the instrumental variables chosen are valid is also investigated. The first requirement of good instruments is that they must be highly correlated with the variable for which they are instrumenting. Table 4 shows that the research and development expenditure and the public spending on education are highly significant in explaining technological innovation.

Table 4. First Stage Regression.

Variable		
Constant term	*0.28**** *(118.01)*	*0.21**** *(17.55)*
Research and development expenditure	*0.14**** *(85.26)*	-
Public spending on education	-	*0.04**** *(14.52)*
R-squared	*0.761*	*0.122*
Number of observations	*2867*	*3782*

[1] The equation estimated in Márquez-Ramos et al. (2007) is considered in this Section:

$$\begin{aligned} \ln X_{ij} &= \alpha_0 + \alpha_1 \cdot \ln Y_i + \alpha_2 \cdot \ln Y_j + \alpha_3 \cdot \ln P_i + \alpha_4 \cdot \ln P_j + \alpha_5 \cdot Adj_{ij} + \alpha_6 \cdot Isl + \alpha_7 \cdot Land + \\ &+ \alpha_8 \cdot CACM + \alpha_9 \cdot CARIC + \alpha_{10} \cdot MERC + \alpha_{11} \cdot NAFTA + \alpha_{12} \cdot CAN + \alpha_{13} \cdot EU + \\ &+ \alpha_{14} \cdot \ln Dist_{ij} + \alpha_{15} \cdot Lang_{ij} + \alpha_{16} \cdot TAI_i + \alpha_{17} \cdot TAI_j + \alpha_{18} \cdot Inf_i + \alpha_{19} \cdot Inf_j + u_{ij} \end{aligned} \quad (3.9)$$

*Notes: ***, **, *, indicate significance at 1%, 5% and 10%, respectively. T-statistics are shown in brackets. The dependent variable is the exporter's technological innovation. The estimation uses White's heteroscedasticity-consistent standard errors.*

The second requirement of good instruments is that they must be uncorrelated with the error term of the export equation. To determine this, the residual of the OLS regression is regressed in the instruments. The results show that the instruments used independently are indeed correlated with the error term. In fact this indicates that the instruments chosen are not the best. However, Cyrus (2002) points out that this test is a very difficult test to pass, and that it may be better to examine the R-squared of these regressions. The results show that the variables used as instruments for income and technology have a low explanatory power (all the instruments have a R-squared value below 0.0063) in the error term regressions.

Finally, the model specification is also checked with both the disaggregated 13-exporter and the aggregated 65-country sample. A model specification error may occur when one or more relevant variables are omitted from the model, or when one or more irrelevant variables are included in the model. Model specification errors can substantially affect the estimated coefficients of regression. The linktest command in STATA is used, and the Ramsey test has been done to test for specification errors.
The linktest is based on the idea that if a regression is properly specified, then it should not be possible to find any additional independent variables that are significant except by chance. The linktest creates two new variables, a variable of the prediction (_hat) and a variable of the square prediction (_hatsq). The model is then refitted using these new variables as predictors. The former should be significant since it is the predicted value, unlike the latter because, if the model is correctly specified, the squared predictions should not have much explanatory power.
In a first step, the linktest is calculated for the disaggregated and aggregated analysis where both the variable of prediction (_hat) and the variable of square prediction (_hatsq) are significant. In a second step, the hypothesis that the model has no omitted variables is rejected with the Ramsey RESET Test. This indicates that further research is needed to improve the specification of the estimated model.

References

- Cyrus, T. L. (2002), "Income in the gravity model of bilateral trade: Does Endogeneity Matter?", *The International Trade Journal* 16(2), 161-180
- Davidson, R. and MacKinnon, J. G. (1993), *Estimation and Inference in Econometrics*, New York, Oxford University Press, Inc.
- Eaton, J. and Kortum, S. (2002), "Technology, geography and trade", *Econometrica* 70(5), 1741-1779.
- Márquez-Ramos, L., Martínez-Zarzoso, I. and Suárez-Burguet, C. (2007), "Technological Innovation, Trade and Development". In: Hakikur Rahman, ed.,

Information and Communication Technologies for Economic and Regional Developments, IGI Publishing: 79-101.