THE ROLE OF DISTANCE IN GRAVITY REGRESSIONS: IS THERE REALLY A MISSING GLOBALISATION PUZZLE?

LAURA MÁRQUEZ-RAMOS, INMACULADA MARTÍNEZ-ZARZOSO AND CELESTINO SUÁREZ-BURGUET*

UNIVERSITAT JAUME I and INSTITUTO DE ECONOMÍA INTERNACIONAL, CASTELLÓN (SPAIN)

IBERO-AMERICA INSTITUTE FOR ECONOMIC RESEARCH, GÖTTIGEN (GERMANY)

Abstract

The main aim of this paper is to investigate the role of distance in gravity regressions in relation to the "missing" globalisation puzzle cited by Coe et al. (2002). Coe et al. (2202) claimed that the non-falling distance coefficients over time is another puzzle that should have been added in Obsfeld and Rogoff (2000) paper, in which six major puzzles in international macroeconomics were cited. In order to do so, a linear and a non-linear gravity model for a cross-section of 65 countries is estimated during the period 1980-1999. Non-linear specifications of the gravity model have been recently used in the literature to address the issue of non-declining transport costs over time (Coe et al., 2002 and Croce et al., 2004). However, we claim that the "missing" globalisation puzzle can be solved with a correct interpretation of the linear specification results. The main findings are that distance has a different effect for developed and developing countries and that the grounds to use a non-linear specification are not clear since the linear specification shows in general terms a better performance than the non-linear.

JEL: F10

---

* Universidad Jaume I and Instituto de Economía Internacional, Castellón, Spain and Ibero-America Institute for Economic Research Universität Göttigen, Germany. Financial support from Fundación Caja Castellón-Bancaja, Generalitat Valenciana and the Spanish Ministry of Education is grateful acknowledged (P1-1B2005-33, Grupos 03-151, INTECO; Research Projects GV04B-030, SEJ 2005-01163 and ACOMP06/047). Address for Correspondence: Universidad Jaume I, Campus del Riu Sec, 12071 Castellón, Spain. E-mail: lmarquez@eco.uji.es. Tel: 0034 964387164. We would like to thank the anonymous referee for their helpful comments and suggestions.

1. Introduction

Several authors have recently focused on the specification of the gravity model. Some of them support the view that non-linear models are preferred to log-linear models. Coe, Subramanian, Tamirisa and Bhavnani (2002) refer to the failure of declining trade costs as an important aspect of globalisation to be reflected in the estimates of the standard gravity model of bilateral trade. They estimate a non-linear specification of the gravity equation and find evidence of the declining importance of geography. Coe et al. (2202) claimed that the non-falling distance coefficients over time in gravity regression is another puzzle that should have been added in Obsfeld and Rogoff (2000) paper, in which six major puzzles in international macroeconomics were cited.[1] Coe et al. (2002) argue that the non-linear specification takes into account zero values for bilateral trade and the level of the estimated distance coefficients is more consistent with the theory. In this line, Croce, Juan-Ramón and Zhu (2004) also estimate a non-linear gravity model and claim that distance has become less relevant over time although the distance coefficient in their yearly regressions declines only slightly (0.06 percentage points in 23 years).

This paper analyses the role of distance in gravity models to test if the "missing" globalisation puzzle really exist. In order to do so, heterogeneity in countries is considered and a log-linear and a non-linear specification of the gravity model are compared to analyse which shows a better performance. The sample includes 65

[1] The six major globalisation puzzles considered by these authors are: the home bias effect in trade, the Feldstein-Horioka saving-investment relationship, the home bias effect in equity portfolios, the international consumption correlations puzzle, the purchasing power parity puzzle and the exchange rate disconnect puzzle.

countries,[2] which account for more than 75% of world trade in 1999, for five selected years during the period 1980-1999.

The theoretical framework is an extension of Helpman and Krugman (1996) model. Their model considers cross-country differences in both, relative factor endowments and relative country size when determining the volume of trade. The extended model considers “hard” and “soft” infrastructure and integration effects (see Suárez-Burguet, Martínez-Zarzoso and Márquez-Ramos, 2005).

The traditional gravity model is augmented with technological variables, since the recent literature has highlighted the importance and significance of these variables as determinants of trade flows (Freund and Weinhold, 2004) and with a remoteness variable (Coca-Castaño, Márquez-Ramos and Martínez-Zarzoso, 2005). A more flexible model that allows for heterogeneity in the slope coefficients according to income levels is also estimated.

The results show that when controlling for heterogeneity in the sample in the linear specification of the gravity model, the declining importance of geography is only found for the adjacency dummy. The distance coefficient decreases 19.35% for developed countries and increases 23.89% for developing countries over the period 1980-1999. However, as claimed by Buch et al. (2004) the magnitude and sign of the distance coefficient is related to the importance of bilateral activities with partners that are far away relative to those that are located nearby. The results from the non-linear specification are less consistent with the theory and present a worse forecast accuracy than the OLS results.

In Section 2 a summary of the methods used to estimate gravity models of trade is discussed. Section 3 describes the variables, discusses the role of distance in gravity

---

[2] 66 countries are considered, however Belgium and Luxembourg data are aggregated.

equations and presents the estimation results of the log-linear specification of the gravity model. A sensitivity analysis is added to test for heterogeneity in the sample. In Section 4 a non-linear specification of the gravity equation is estimated. In Section 5, the log-linear and non-linear models are compared. Finally, Section 6 concludes.

2. The gravity model of bilateral trade

The first authors to apply the gravity model to international trade flows were Tinbergen (1962) and Pöyhönen (1963). This model, in its basic form, assumes that trade between countries can be compared to the gravitational force between two objects: it is directly related to countries' size and inversely related to the distance between them. Theoretical support for the research in this field was originally very poor but since the second half of the 1970s several theoretical developments have appeared in support of the gravity model. Anderson (1979) made the first formal attempt to derive the gravity equation from a model that assumed product differentiation. Bergstrand (1985, 1989) also explored the theoretical determination of bilateral trade in two papers, in which gravity equations were associated with simple monopolistic competition models. Helpman and Krugman (1996) used a differentiated product framework with increasing returns to scale to justify the gravity model. Deardorff (1995) has proven that the gravity equation characterises many models and can be justified from standard trade theories.

According to the generalised gravity model of trade, the volume of exports between pairs of countries, X_{ij}, is a function of their incomes (GDPs), their populations, their geographical distance and a set of dummies. Technology could alter the effect of geographical barriers on trade, but, as far as the missing globalisation puzzle is concerned, the usual gravity model estimation has not found that trade becomes less sensitive to distance over time.

The differences in the abovementioned theories help to explain the various specifications and some diversity in the results of the empirical applications.

In this paper, the theoretical framework is taken from Suárez-Burguet et al. (2005). These authors analyse the role played by cross-country differences in both, relative factor endowments and relative country size when determining the volume of trade. They developed a model based on Helpman and Krugman (1996) and introduced trade barriers, and "hard" and "soft" investment in infrastructure as determinants of the volume of trade. The model supports the notion that comparative advantage determines international trade, adding factors with a positive influence on production factors: "hard" (transport infrastructure) and "soft" (technological innovation) investment in infrastructure. Although the gains of increasing trade are attenuated by the resistance imposed by geographical barriers, infrastructure endowment also determines countries' specialisation and trade flows.

3. Augmented log-linear gravity model

3.1. Data sources and variables

Table A.1 in Appendix[3] shows a summary of the data used in our analysis. Some additional explanations are needed with respect to technological variables. The *ArCo Technology Index* is used in this paper to measure technological innovation (Archibugi and Coco, 2002). The ArCo is more complete compared to other indices, since it takes into account a wider array of variables related to technological innovation.[4] We use data

---

[3] Appendix A. Table A.1. The first column lists the variables used for empirical analysis; the second column outlines a description of the variables, and the third column shows the data sources.

[4] This index takes into account three dimensions: Creation of technology (number of patents, number of scientific papers), diffusion of technology (Internet penetration, telephone penetration, electricity consumption) and development of human skills (gross tertiary science and engineering enrolment, mean years of schooling, adult literacy rate).

of ArCo for 1990 and 1999. We use the ArCo of 1990 as a measure of technological innovation in regressions for 1980, 1985 and 1990, and the ArCo of 1999 as a measure of technological innovation in regressions for 1995 and 1999.

3.2. The role of distance in gravity models

Gravity models of trade use distance between countries as a proxy for transport costs, assuming that transport cost from the exporter country to the importer country are the same that transport costs from the importer to the exporter country ($t_{ij}=t_{ji}$). However, Anderson and van Wincoop (2004) emphasise the need to obtain better transport cost measures and to use these measures to expand gravity models and deal with the endogeneity of the transport cost variable in gravity equations. Geographical distance may represent a series of factors such as cultural proximity, a shared history, a perception of closeness and information costs rather than acting as a proxy for transport costs, since they tend to be fixed according to the supply and demand conditions in the market (Márquez-Ramos, Martínez-Zarzoso, Pérez-García and Wilmsmeier, 2006). Therefore, it is still not clear what is the real information embodied in the distance coefficient estimated in gravity regressions. In spite of these limitations, distance variables are commonly used in gravity regressions as a proxy for transport costs with the implicit assumption that distance costs are a linear function of distance.

Buch, Kleinert and Toubal (2004) show that if distance costs decrease proportionally for all countries, no change should be found in the distance coefficient, since the distance coefficient measures the relative importance of economic relationships between trading partners located far away, as opposed to those located nearby. An increase in the constant term would indicate that the distance costs decrease, however when applied to real data, this effect is mixed with an omitted variables effect also included in the constant term. Buch et al. (2004) distinguish three scenarios. First, when the distance costs decrease proportionally for all countries all the information about the positive effect of decreasing distance costs is included in the constant term, which increases in magnitude. Second, when the distance costs decrease non-proportionally and the decrease is greater for smaller distances the distance coefficient increases over time. Third, when the distance costs decrease non-proportionally and the decrease is smaller for smaller distances the distance coefficient decreases over time in absolute terms. In other words, a decrease in the distance coefficient indicates that trade with countries that are far away from the home country increases relative to trade with countries that are closer to the home country, whereas an increase indicates that trade with countries that are far away decreases relative to trade with countries that are closer to the home country.

In what follows the interpretation of Buch et al. (2004) is used to explain the results obtained in gravity regressions. Since developed and developing countries have a different behaviour in the context of gravity equations (see Márquez-Ramos and Martínez-Zarzoso, 2005), the evolution of the distance coefficient is also analysed considering heterogeneity in the sample.

3.3. Estimation results

65-country sample

We start by estimating a log-linear version of the gravity model (Bergstrand, 1985, 1989; Deardorff, 1995). The model is augmented with technological innovation and remoteness, since recent literature in gravity models of international trade has highlighted that theoretically these models are determined by relative trade barriers and not only by absolute trade barriers between the exporter and the importer country. A number of dummies representing integration, geographical and cultural characteristics are also added in order to analyse the impact of these factors on international trade. The model is expressed in additive form using a logarithmic transformation. The estimated equation is:

$$\begin{aligned}\ln Xij = {} & \alpha_0 + \alpha_1 \cdot \ln Y_i + \alpha_2 \cdot \ln Y_j + \alpha_3 \cdot \ln P_i + \alpha_4 \cdot \ln P_j + \alpha_5 \cdot Adj_{ij} + \alpha_6 \cdot Isl_i + \alpha_7 \cdot Isl_j + \\ & + \alpha_8 \cdot Land_i + \alpha_9 \cdot Land_j + \alpha_{10} \cdot \ln rem_i + \alpha_{11} \cdot \ln rem_j + \alpha_{12} \cdot CACM + \alpha_{13} \cdot CARIC + \\ & + \alpha_{14} \cdot MERC + \alpha_{15} \cdot NAFTA + \alpha_{16} \cdot CAN + \alpha_{17} \cdot UE + \alpha_{18} \cdot \ln Dist_{ij} + \alpha_{19} \cdot Lang_{ij} + \\ & + \alpha_{20} \cdot ArCo_i + \alpha_{21} \cdot ArCo_j + u_{ij}\end{aligned} \qquad (1)$$

where *ln* denotes natural logarithms.

The model is estimated with data for 65 countries in 1980, 1985, 1990, 1995 and 1999 and a total of 4160 (65*64) bilateral trade flows are obtained (Appendix, Figure 1). The presence of missing/zero values in the bilateral trade flows data for different years slightly reduces the sample (e.g. in 1999 to 3347 observations). We perform OLS estimation on the double log specification as given by equation (1).

X_{ij} denotes the value of exports from the exporter country i to the importer country j, Y_i and P_i are income and population in the exporter's market, Y_j and P_j are income and population in the destination market, Adj_{ij} is a dummy that takes a value of 1 when countries share the same border and zero otherwise, Isl_i and Isl_j take a value of 1 when the exporter or the importer are islands, $Land_i$ and $Land_j$ are dummies that take the value of 1 when the exporter and the importer are *landlocked* countries. Rem_i and Rem_i measure the level of remoteness in the exporter and the importer country. These variables are calculated according to equation (2) and (3), as in Coca-Castaño et al. (2005). It is intended to measure the average distance of an exporter country from all its trading partners,

$$\mathrm{Re}\boldsymbol{m}_i = \sum_j \left(\frac{\boldsymbol{Y}_j}{\boldsymbol{Y}^w} \right) \boldsymbol{Dist}_{ij} \qquad (2)$$

$$\mathrm{Re}\, m_j = \sum_i \left(\frac{Y_i}{Y^w} \right) Dist_{ij} \qquad (3)$$

where Y^w represents the world income. Table A.2 shows the values of the remoteness variable for the years analysed.

CACM is a dummy that takes a value of 1 when both countries belong to the Central American Common Market, CARIC is a dummy that takes a value of 1 when both countries belong to the Caribbean Community, MERC is a dummy that takes a value of 1 when both countries belong to Mercosur, NAFTA takes a value of 1 when countries are members of the North American Free Trade Area, CAN is a dummy representing Andean Nations Community members and UE takes a value of 1 when countries are members of the European Union. For comparison purposes the integration dummies are included over the entire period. In this way we have the same variables in all regressions. Frankel (1997) justifies the inclusion of FTA dummies even before a formal trading bloc has come into being; indicating that informal ties between the countries usually existed previously.

Since suitable direct measures of trade costs are unavailable, geographical distance between countries is often used as a proxy for transport costs in gravity equations, so $Dist_{ij}$ is the geographical great circle distance in kilometres between the capitals of country *i* and *j*. $Lang_{ij}$ is a dummy for countries sharing the same language, this variable captures generally information costs in countries. $ArCo_i$ and $ArCo_j$ represent the effect of technological innovation. These variables are expected to have a positive coefficient according to the theoretical framework in Suárez-Burguet et al. (2005).[5] Finally, $\boldsymbol{u}_{ij}$ is assumed to be independently and identically distributed among countries.

Table 1 shows estimation results for equation (1) for different years. Income variables are significant and have the expected positive sign. Population variables are negative

[5] Suárez-Burguet et al. (2005) introduce the effect of technological innovation on trade by assuming that it reduces final prices. Then, an increase in these variables will raise trade volumes.

indicating the existence of an absorption effect in the 65-country sample. Martínez-Zarzoso and Nowak-Lehmann (2003) point out that the coefficient of population can be negative or positive signed, depending on whether the country exports less when it is big (absorption effect) or whether a big country exports more than a small country (economies of scale).

The importance of the adjacency dummy decreases over time, starting in 1980 with an estimated coefficient of 0.46, and ending in 1999 with a coefficient of 0.31. Landlockedness seems to have a higher negative effect on trade nowadays than in the past. Exporter's remoteness is significant and shows the expected positive sign whereas importer's remoteness is significant and positive signed in 1980, 1995 and 1999 and negative and non-significant in 1985 and 1990. A positive sign is also found in other studies of trade such as Coca-Castaño et al. (2005).

Table 1 also shows the increasing importance of most Latin America agreements (CACM, CARICOM, MERCOSUR and CAN) on trade flows, whereas the NAFTA dummy is not significant and the European Union (UE dummy) is only significant in 1980 and 1990. A possible explanation for this result may be the existence of heterogeneity in the sample (see Márquez-Ramos and Martínez-Zarzoso, 2005).

Distance shows its expected puzzling result, a slightly increasing negative impact on trade flows. The coefficients obtained for the language dummy increases over time, indicating that easier verbal communication between trading partners facilitates trade to a higher extent in the 90s than in the 80s. We find that this model has a high explanatory power in terms of the R^2.

Sensitivity analysis

Márquez-Ramos and Martínez-Zarzoso (2005) show that the pooling assumption is rejected in a sample of countries with different levels of economic development, since the determinants of trade have different coefficients for high and low-income countries. Following Márquez-Ramos and Martínez-Zarzoso (2005) two groups of countries are considered. We use DP dummy which takes the value of one when the countries are richer than the average in the sample and zero otherwise. Results of the interaction dummies (Table 2) show the existence of heterogeneity in the sample. The distance coefficient is statistically different for developed and developing countries. The positive signed coefficients when interacting with DP dummy mean that these variables are more important to foster trade in developing countries than in developed ones.

Other geographical variables do not show much evidence of heterogeneity. The adjacency coefficient decreases over time, and has a similar magnitude for developed and developing countries from 1985 onwards. The coefficient of the landlockedness variable increases over time and significantly differs for developed and developing countries only in 1999. Language has a higher effect on trade flows in developing than in developed countries from 1985 onwards. Coefficient heterogeneity is also observed in the technological innovation variables, which show higher coefficients for developing countries.

In relation to the income coefficients, Garman, Petersen and Gilliard (1998) analyse economic integration in a number of developing countries and support the notion that the costs and benefits of integration are unevenly distributed among members of an integration agreement in favour of the richest countries. They find that the income coefficients obtained for Latin American countries have a smaller magnitude than those reported in other studies of European trade. Our results partially support this evidence since the exporter income coefficients are higher for high-income economies from 1990 (the interaction dummies take a positive value).

Going back to distance coefficients, we observe that the distance coefficient is considerably lower for developed countries and from the year 1985 onwards the differences are statistically significant with respect to developing countries. The magnitude of the coefficient shows an increase for both groups of countries from 1990 onwards. For the whole period 1980-1999, a 19.35% decrease in the distance coefficient is observed for the developed countries, whereas a 23.89% increase in the distance coefficient is observed for developing countries. According to Buch et al. (2004) changes in the distance coefficient over time[6] cannot be interpreted in terms of rising or falling distance costs. Instead, this coefficient measures how important bilateral economic activities with partners that are far away are relative to those with partners that are close to the home country. These authors show that the distance coefficient might remain unchanged and that changing distance costs is reflected in the constant term. In this line, the evolution of the constant term estimated in Tables 1 and 2 and the evolution of the distance coefficient for developed and developing countries are analysed. Figures 2 and 3 show the evolution over time according to the results of the OLS regressions.

The constant term increases in absolute value over time (Figure 2) when the 65 countries are considered (Table 1) and also when heterogeneity is considered (Table 2). This effect could be a sign of a reduction in transport costs over time; however this effect is mixed with an omitted variables effect which might also change the value of the constant term.

With respect to the distance coefficient, the different scenarios considered in Buch et al. (2004) are taken into account to classify each group of countries. The developing

[6] When the distance coefficient is obtained from the estimation of gravity equations using OLS at different points in time.

countries can be placed in the scenario 2, since the magnitude of the distance coefficient increases over the period 1980-1999, whereas the developed countries can be placed in scenario 3, since the magnitude of the distance coefficient decreases over the period 1980-1999. For developing countries, export flows for small distances increase over time, whereas export flows for large distances decrease over time, therefore trade with countries far away decreases relative to trade with countries nearby. The opposite applies to developed countries.

4. Estimation of a non-linear model

Coe et al. (2002) estimate a non-linear specification of the gravity equation and find evidence for the declining importance of geography. They prefer the non-linear specification since it takes into account zero values for bilateral trade and the level of the estimated distance coefficients is more consistent with theory. Moreover they point out that "the non-linear models do a much better job of explaining bilateral trade than do the log-linear models".[7] In a more recent paper, Croce et al. (2004) study the performance of four trading blocs (MERCOSUR, NAFTA, CACM and the Andean Community). They estimate a non-linear gravity equation by NLS for each year during the period 1978-2001 and find that it explains an important part of the determinants of trade flows: income, population, distance, adjacency, language and integration variables have the expected sign and are significant. These authors state that "in spite of the technological progress experienced in the last two decades, global economic geography is still relevant. This represents one of the puzzles of globalization".[8] Both papers find that distance has become less relevant over time by using non-linear estimation techniques. The estimated coefficients fell over time, but their values still reveal the

[7] Coe et al. (2002), page 21.

[8] Croce et al. (2004), page 9.

importance of geography on trade. Although Coe et al. (2004) mention in page 16 that they also run regressions excluding developing countries from the sample, they do it to test the robustness of their results to potential data weaknesses and not to test for heterogeneity in the sample. They say that the results, which are not reported, were "very similar" to those based on the full sample.

In order to compare log-linear and non-linear specification, we derive the empirical model to be estimated, which is given by equation (4):

$$
\begin{aligned}
X_{ij} &= \left(P_i^{c(1)} \cdot ArCo_i^{c(2)} \cdot Y_j^{c(3)} + P_j^{c(4)} \cdot ArCo_j^{c(5)} \cdot Y_i^{c(6)}\right) \cdot \frac{rem_i^{c(8)} \cdot rem_j^{c(9)} \cdot e^{\mu ij}}{Dist_{ij}^{c(7)}} + \varepsilon_{ij} \\
\mu_{ij} &= \alpha_1 \cdot CACM + \alpha_2 \cdot CARIC + \alpha_3 \cdot MERC + \alpha_4 \cdot NAFTA + \alpha_5 \cdot CAN + \\
&+ \alpha_6 \cdot UE + \alpha_7 \cdot Adj_{ij} + \alpha_8 \cdot Lang_{ij} + \alpha_9 \cdot Land_i + \alpha_{10} \cdot Land_j + \alpha_{11} \cdot Isl_i + \alpha_{12} \cdot Isl_j
\end{aligned}
\tag{4}
$$

where bilateral export flows between countries is the dependent variable, and the explanatory variables are the same included in the log-linear model. Variables representing integration agreements, adjacency, language, landlocked and island dummies are included in the non-linear model by means of an exponential function, as is done in Croce et al (2004).

The optimisation process to estimate a non-linear model has three main parts: obtaining the starting parameter values, updating the candidate parameter vector at each interaction, and determining when we have reached the optimum. If the objective function is globally concave, there is a single maximum and any algorithm that improves the parameter vector at each iteration will eventually find this maximum. If the objective function is not globally concave, different algorithms may find different local maxima. However, all iterative algorithms suffer from the same problem of being unable to distinguish between local and global maxima. Eviews uses the Marquardt algorithm to solve by NLS. This algorithm modifies the Gauss-Newton algorithm by adding a correction matrix to the Hessian approximation.

Iterative estimation procedures require starting values for the coefficients of the model and the closer to the true values they are, the better. However, there are no general rules for selecting them. We have used the simple average of the values obtained in Croce et al. (2004) and Coe et al. (2002) for the starting values of income, population and distance.[9] For the CAN, CACM, MERC and NAFTA dummies, adjacency and language variables, we include the coefficients obtained in Croce et al. (2004) as starting values. For remoteness we have used the values of the coefficient obtained in Coe et al. (2002). Finally, the OLS coefficients are used as starting values for the rest of explanatory variables.

The estimation process achieves convergence if the maximum change in the coefficients is below the specified value. We have used a convergence criterion of 0.001. When a non-linear model converges, the standard statistical results and tests are asymptotically valid.

Table 3 shows final results and allows comparing non-linear estimations with linear estimations obtained in Table 1. Income variables are significant and have the expected positive sign. Population variables are positive indicating the existence of scale economies in the 65-country sample, this result is opposite to the obtained in log-linear estimation. The magnitude of the adjacency coefficient decreases over time in the 90s, although in 1999 is still relevant. Unexpected sings are found for island and remoteness. Decreasing evidence of the importance of being an island is found. Remoteness has a negative sign indicating that more remote countries trade less; the opposite (expected) sign was obtained in log-linear regressions. Landlocked has a lower positive effect (unexpected) on trade in 1999 than in 1980 for the importer country. Table 3 shows a

[9] Distance is included with a positive sign because, in the non-linear specification estimated in this paper, it enters in the denominator.

decreasing positive effect of two Latin American agreements (CACM and CAN) on trade flows, and an increasing positive effect of CARICOM, Mercosur and NAFTA agreements. According to the non-linear results, the European Union agreement (UE dummy) has a decreasing positive effect on trade flows. Finally, a greater technological innovation endowment in both the exporter and the importer country leads to higher international trade flows. The coefficient for the exporter's technological innovation decreases over time, whereas it increases for the importer. This result is also different to the one obtained in log-linear regressions, where the coefficients of both variables decrease over time, and the importers' technological innovation has a lower effect on trade than the exporters' technological innovation.

Table A.3 in Appendix shows the different coefficients obtained in the present paper, in Coe et al. (2002) and in Croce et al. (2004). These coefficients have similar elasticities in the three papers for income, adjacency and some integration agreements, although the units of measure of the dependent variable are different in Coe et al. (2002).

5. Log-linear and non-linear model comparison

When non-linear and log-linear models are compared, the adjusted R-squared is not a good indicator since it depends on the selected dependent variable and different scales are used in both specifications.[10] Coe et al. (2002) use the exponentials of the predicted values in the log-linear regression to calculate an adjusted R-squared transformed to be comparable to the adjusted R-squared in the non-linear model, however the non-linear specification does not include a constant term and this should be considered for the purpose of comparing models.

In this paper we compare the forecast accuracy between log-linear and non-linear models. Henderson and Millimet (2006) compare different specifications of the gravity

[10] In the log-linear model the dependent variable is lnX, whereas in the non-linear is X.

model: parametric (linear and non-linear) and non-parametric. The authors compute four measures of forecast accuracy: the squared correlation coefficient between the actual value and the predicted value, the predicted mean squared error, the predicted mean absolute error and the predicted mean absolute percentage error. Their result indicate that, independently of the accuracy measure used, the parametric estimations are preferred to the non-parametric and within the parametric models the log-linear models present a better forecasting accuracy than the non-linear models.

In the present paper, the predicted root mean squared error, the predicted mean absolute error and the predicted mean absolute percentage error are chosen to compare log-linear and non-linear models. According to these forecasting measures, the log-linear model outperforms the non-linear model and the non-linear model underpredicts the volume of trade in our sample.

Another way to compare both specifications is to use the estimation results to generate predictions in order to compare the forecast evaluation of the linear and non-linear models. Stavins and Jaffe (1990) use a goodness of fit statistic equal to one minus Theil's U-statistic based on comparing predicted and actual values for the dependent variable (S&J goodness of fit). The Theil inequality coefficient lies between 0 and 1 and a value of zero indicates a perfect fit. We can compare log-linear and non-linear models with this measure since it is scale invariant. The S&J (1990) goodness of fit values are shown at the bottom of Tables 1, 2 and 3. The values obtained are always higher in log-linear models, indicating that these models give a better forecast accuracy than non-linear models..

Finally, both models (Table 1 and 3) are also compared with the Akaike Info Criterion (AIC). In order to do so we transform the AIC obtained in log-linear specifications since in the log-linear model the dependent variable is the natural logarithm of exports whereas in the non-linear is exports.[11] Smaller values of the AIC indicate that a model is preferred. According to this criterion, the linear gravity model is a better alternative.

Overall, the results in this paper indicate that the linear gravity model is still a very good technique to estimate gravity equations and, although one of the main criticisms against using linear gravity equations to estimate trade flows is the failure to reflect declining trade costs,[12] we argue that evidence of declining trade costs can also be shown in linear gravity estimations, when heterogeneity in the sample is considered and when the results are correctly interpreted. Therefore, it is not clear why a non-linear estimation should resolve this puzzle in a better way than a linear estimation. Opposite effects are found for a number of variables in the non-linear estimations and the log-linear model outperforms the non-linear in terms of accuracy and AIC criteria.

6. Conclusions

In this paper, we estimate a log-linear and a non-linear gravity equation in order to compare the evolution over time of the determinants of bilateral trade flows. Geographical (distance, adjacency, being an island and landlockedness), social (integration agreements and common language) and technological innovation (ArCo index) are included in the model. Since the literature in gravity equations has highlighted the importance of relative trade barriers, remoteness variables are also included.

To control for heterogeneity effects, we divide the countries of our sample according to their level of economic development in two groups: high-income and low-income countries. Geographical and social factors influence trade in both groups, but the effects are more pronounced for low-income than for the high-income countries in distance and language. Technological endowments also seem to have a higher effect on trade in

[11] $\boldsymbol{AIC}_{\log} = \boldsymbol{AIC} + 2 \cdot \overline{\ln \boldsymbol{Y}}$

[12] This is considered puzzling since transport costs have declined over time.

developing economies. These variables are therefore of great importance to increase the participation of the poorest economies in the world economy.

Concerning distance, we find that the coefficient of this variable increases over time for developing countries and it decreases over time and it is much lower in magnitude for developed countries. Following Buch et al. (2004) interpretation of the evolution of the constant term over time, the observed increase in the constant term from 1995 onwards, which is more pronounced when the heterogeneity in the sample is considered, points towards a decreasing importance of trade costs.

A non-linear specification of the main model is estimated and compared with the log-linear one. The distance coefficient is clearly lower than the one found in the linear estimations and decreases over time, however, opposite unexpected effects are found for a number of variables in the non-linear estimations. Additionally, in terms of accuracy and AIC criteria the log-linear outperforms the non-linear model. Although the non-linear specification of the gravity equation seems to be a good approach to analyse the determinants of international trade given its high value of the R-squared, the log-linear specification of the gravity model shows a better forecast accuracy on trade flows and a lower AIC value.

REFERENCES

Anderson, J. E. (1979), "A theoretical foundation for the gravity equation", *American Economic Review,* 69, 106-116.

Anderson, J. E. and van Wincoop, E. (2004), "Trade Costs", *Journal of Economic Literature*, 42, 691-751.

Archibugi, D. and Coco, A. (2002), "A new indicator of technological capabilities for developed and developing countries (ArCo)", *Italian National Research Council.*

Bergstrand, J. H. (1985), "The gravity equation in international trade: Some microeconomic foundations and empirical evidence", *The Review of Economics and Statistics*, 67, 474-481.

Bergstrand, J. H. (1989), "The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade", *The Review of Economics and Statistics*, 71, 143-153.

Buch, C. M., Kleinert, J. and Toubal, F. (2004), "The distance puzzle: on the interpretation of the distance coefficient in gravity equations", *Economics Letters*, 83, 293-298.

Central Intelligence Agency – CIA– (2003), *The World Factbook.* http://www.odci.gov/cia/publications/factbook

Coca-Castaño, P., Márquez-Ramos, L. and Martínez-Zarzoso, I. (2005), "Infraestructuras, costos de transporte y flujos de comercio", *Revista de Análisis Económico*, 20(1), 3-22.

Coe, D. T., Subramanian, A., Tamirisa, N. T. and Bhavnani, R. (2002), "The missing globalization puzzle", IMF Working Paper No. 02/171.

Croce, E., Juan-Ramón, V. H. and Zhu, F. (2004), "Performance of Western Hemisphere Trading Blocs: A Cost-Corrected Gravity Approach", IMF Working Paper No. 04/109.

Davidson, R. and MacKinnon, J. G. (1993), *Estimation and Inference in Econometrics*, Oxford University Press.

Deardorff, A. V. (1995), "Determinants of bilateral trade: Does gravity work in a Neo-classical word", NBER Working Paper 5377.

Eaton, J. and Kortum, S. (2002), "Technology, geography and trade", *Econometrica*, Vol. 70, Nº 5, 1741-1779.

Frankel, J. (1997), *Regional Trading Blocs in the Wold Economic System,* Washington, DC: Institute for International Economics.

Freund, C. L. and Weinhold, D. (2004), "The effect of the Internet on international trade", *Journal of International Economics*, 62, 171-189.

Garman, G., Petersen J., and Gilliard, D. (1998), "Economic integration in the Americas: 1975-1992", *Journal of Applied Business Research*, Laramine, Summer.

Great circle distances between cities (2003): http://www.wcrl.ars.usda.gov/cec/java/lat-long.htm

Helpman, E. and Krugman, P. R. (1996), *Market Structure and Foreign Trade. Increasing Returns, Imperfect Competition, and the International Economy*, Cambridge, MA: MIT Press.

Henderson, D. J. and Millimet, D. L. (2006), "Is gravity linear?", Working Paper Series 2006: 0603, Department of Economics at Binghamton University, New York.

Márquez-Ramos, L. and Martínez-Zarzoso, I. (2005), "Does heterogeneity matter in the context of the gravity model?", *Economics Bulletin,* 6(17), 1-7.

Márquez-Ramos, L., Martínez-Zarzoso, I., Pérez-García, E. and Wilmsmeier, G. (2006), "Factores Determinantes del Flete: Análisis para el Caso Español". In Actas del VII Congreso de Ingeniería del Transporte. 1ª Edición. ISBN: 84-689-8340-3. Eds: Menéndez-Martínez, J. M., Guirao-Abad, B. and Rivas-Álvarez, A. ANGAMA Artes Gráficas.

Martínez-Zarzoso, I. and Nowak-Lehmann, F. (2003), "Augmented Gravity Model: An empirical application to Mercosur-European Union trade flows", *Journal of Applied Economics*, vol. VI, 2, 291-316.

Obstfeld, M., and K. Rogoff (2000), "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?", NBER Working Paper Series, Working Paper 7777.

Pöyhönen, P. (1963), "A tentative model for the volume of trade between countries", *Weltwirtschaftliches Archiv* 90, 93-99.

Statistics Canada (2001), *World Trade Analyzer*. http://www.statcan.ca/

Stavins, R. N. and Jaffe, A. B. (1990), "Unintended Impacts of Public Investments on Private Decisions: The Depletion of Forested Wetlands", *American Economic Review*, 80, 337-352.

Suárez-Burguet, C., Martínez-Zarzoso, I. and Márquez-Ramos, L. (2005). "The non-linear specification of the gravity model: An empirical application on international trade". PD-ECO 2005/1. Castellón, Spain, Universitat Jaume I.

Tinbergen, J. (1962), "Shaping the world economy. Suggestions for an international economic policy", The Twentieth Century Fund, New York.

World Bank (2001), *World Development Indicators*, Washington.

TABLES

Table 1. Determinants of international trade. Augmented log-linear gravity model, OLS estimation.

	1980	1985	1990	1995	1999
Constant term	-44.83***	-43.13***	-46.24***	-50.94***	-49.96***
	(-15.40)	(-14.33)	(-17.53)	(-22.80)	(-23.40)
Exporter's income	1.12***	1.53***	1.49***	1.52***	1.46***
	(13.87)	(16.74)	(18.00)	(23.46)	(22.69)
Importer's income	0.93***	0.84***	1.12***	1.23***	1.21***
	(10.81)	(9.18)	(13.87)	(18.81)	(19.02)
Exporter's population	-0.23***	-0.66***	-0.55***	-0.52***	-0.40***
	(-2.88)	(-7.03)	(-6.68)	(-7.71)	(-5.99)
Importer's population	-0.17**	-0.04	-0.33***	-0.41***	-0.42***
	(-2.01)	(-0.42)	(-4.08)	(-6.14)	(-6.44)
Adjacency dummy	0.46**	0.40**	0.58***	0.37***	0.31**
	(2.41)	(2.36)	(3.44)	(2.92)	(2.24)
Exporter island dummy	-0.05	-0.25**	-0.12	-0.30***	-0.22***
	(-0.40)	(-2.15)	(-1.30)	(-3.52)	(-2.63)
Importer island dummy	-0.08	0.18	0.20**	-0.12	0.06
	(-0.72)	(1.63)	(2.33)	(-1.58)	(0.75)
Exporter landlocked dummy	-0.25**	-0.26**	-0.15	-0.47***	-0.42***
	(-2.08)	(-2.18)	(-1.35)	(-5.16)	(-4.85)
Importer landlocked dummy	-0.44***	-0.65***	-0.43***	-0.46***	-0.72***
	(-3.38)	(-5.17)	(-3.63)	(-4.87)	(-7.61)
Exporter's remoteness	1.53***	1.61***	1.47***	1.09***	1.19***
	(7.02)	(7.17)	(7.70)	(7.09)	(8.19)
Importer's remoteness	0.37*	-0.25	-0.23	0.43***	0.29*
	(1.66)	(-1.13)	(-1.19)	(2.75)	(1.84)
CACM dummy	-	-	1.49***	1.69***	1.98***
			(5.64)	(7.01)	(9.03)
CARICOM dummy	4.69***	4.17***	4.41***	5.04***	4.89***
	(25.76)	(23.25)	(25.77)	(6.27)	(5.37)
MERCOSUR dummy	-	-	-	0.61	0.73*
				(1.56)	(1.86)
NAFTA dummy	-	-	-	-0.43	-0.17
				(-0.79)	(-0.28)
CAN dummy	-	-	0.41*	1.22***	1.50***
			(1.89)	(6.26)	(5.70)
UE dummy	0.45***	0.19	0.46***	0.10	-0.02
	(2.99)	(1.20)	(4.30)	(1.29)	(-0.20)
Distance	-1.03***	-1.08***	-0.98***	-1.08***	-1.18***
	(-18.16)	(-19.96)	(-20.83)	(-26.07)	(-28.47)
Language dummy	0.41***	0.35***	0.56***	0.65***	0.68***
	(3.80)	(3.18)	(5.50)	(7.30)	(7.84)
Exporter's ArCo	4.61***	3.00***	3.33***	3.18***	3.49***
	(10.47)	(6.25)	(7.27)	(10.60)	(12.10)
Importer's ArCo	2.51***	3.02***	1.94***	1.54***	1.46***
	(5.42)	(6.04)	(4.38)	(4.74)	(4.59)
R-squared	0.65	0.66	0.70	0.78	0.79
Number of observations	2440	2408	2926	3334	3347
Akaike Info Criterion	22.91	22.88	23.77	24.52	24.55
Root Mean Squared Error	1.76	1.75	1.69	1.52	1.49
Mean Absolute Error	1.33	1.31	1.26	1.13	1.11

Mean Absolute Percentage Error	20.9	20.73	18.66	15.69	15.52
S&J (1990) goodness of fit	0.91	0.91	0.92	0.93	0.93

Notes: ***, **, * indicate significance at 1%, 5% and 10%, respectively. T-statistics are in brackets. The dependent variable is the natural logarithm of exports in value (current US$). Income, population and distance are also in natural logarithms. The estimation uses White's heteroscedasticity-consistent standard errors.

Table 2. Determinants of international trade. Augmented log-linear gravity model with country heterogeneity.

	1980	1985	1990	1995	1999
Constant term	-40.37***	-35.93***	-45.82***	-49.05***	-46.68***
	(-12.08)	(-10.24)	(-15.09)	(-19.51)	(-18.97)
Exporter's income	1.11***	1.49***	1.32***	1.33***	1.23***
	(10.83)	(12.73)	(12.39)	(15.93)	(15.13)
DP*Exporter's income	**-0.48****	**-0.64*****	**0.36***	**0.28***	**0.33****
	(-2.26)	**(-2.77)**	**(1.80)**	**(1.76)**	**(2.02)**
Importer's income	0.82***	0.69***	0.96***	1.08***	1.07***
	(7.41)	(5.79)	(9.31)	(12.19)	(12.58)
DP*Importer's income	**0.24**	**0.42***	**0.82*****	**0.53*****	**0.29***
	(1.04)	**(1.77)**	**(4.01)**	**(3.35)**	**(1.82)**
Exporter's population	-0.23**	-0.61***	-0.33***	-0.25***	-0.08
	(-2.30)	(-5.05)	(-3.09)	(-2.89)	(-0.95)
DP*Exporter's population	**0.55****	**0.67*****	**-0.50****	**-0.49*****	**-0.58*****
	(2.39)	**(2.70)**	**(-2.38)**	**(-2.84)**	**(-3.39)**
Importer's population	-0.06	0.12	-0.16	-0.25***	-0.29***
	(-0.54)	(0.98)	(-1.49)	(-2.67)	(-3.17)
DP*Importer's population	**-0.23**	**-0.40**	**-0.87*****	**-0.58*****	**-0.28**
	(-0.94)	**(-1.58)**	**(-4.03)**	**(-3.49)**	**(-1.64)**
Adjacency dummy	0.72***	0.64***	0.70***	0.61***	0.56***
	(2.85)	(2.97)	(3.07)	(3.24)	(2.76)
DP*Adjacency dummy	**-0.66****	**-0.43**	**-0.07**	**-0.26**	**-0.28**
	(-2.08)	**(-1.54)**	**(-0.26)**	**(-1.19)**	**(-1.21)**
Exporter island dummy	-0.06	-0.37**	-0.15	-0.28**	-0.15
	(-0.35)	(-2.23)	(-1.14)	(-2.29)	(-1.24)
DP*Exporter island dummy	**0.18**	**0.59*****	**0.15**	**0.05**	**-0.05**
	(0.79)	**(2.67)**	**(0.80)**	**(0.33)**	**(-0.29)**
Importer island dummy	-0.13	0.13	0.13	-0.20*	0.08
	(-0.85)	(0.81)	(1.07)	(-1.83)	(0.82)
DP*Importer island dummy	**0.15**	**0.17**	**0.10**	**0.17**	**-0.02**
	(0.69)	**(0.78)**	**(0.59)**	**(1.13)**	**(-0.11)**
Exporter landlocked dummy	-0.21	-0.21	-0.16	-0.49***	-0.44***
	(-1.33)	(-1.31)	(-1.06)	(-3.95)	(-3.76)
DP*Exporter landlocked dummy	**0.11**	**0.18**	**0.14**	**0.22**	**0.26***
	(0.54)	**(0.87)**	**(0.72)**	**(1.44)**	**(1.75)**
Importer landlocked dummy	-0.42**	-0.61***	-0.49***	-0.43***	-0.73***
	(-2.45)	(-3.69)	(-3.21)	(-3.37)	(-5.73)
DP*Importer landlocked dummy	**0.16**	**0.21**	**0.31**	**0.09**	**0.27**
	(0.67)	**(0.91)**	**(1.49)**	**(0.54)**	**(1.56)**
Exporter's remoteness	1.49***	1.67***	1.86***	1.30***	1.35***
	(5.16)	(5.47)	(7.37)	(6.67)	(7.24)
DP*Exporter's remoteness	**-0.39**	**-1.06*****	**-1.28*****	**-0.75*****	**-0.68*****
	(-1.22)	**(-3.07)**	**(-4.40)**	**(-3.43)**	**(-3.17)**
Importer's remoteness	0.02	-0.95***	-0.47*	0.20	-0.01
	(0.08)	(-3.08)	(-1.86)	(1.01)	(-0.06)
DP*Importer's remoteness	**0.44**	**1.03*****	**0.36**	**0.28**	**0.30**

	(1.38)	**(3.09)**	**(1.27)**	**(1.29)**	**(1.33)**
CACM dummy	-	-	1.21***	1.34***	1.59***
			(3.95)	(4.66)	(6.26)
CARICOM dummy	4.66***	4.21***	4.53***	5.13***	4.81***
	(20.21)	(17.85)	(22.38)	(6.25)	(5.09)
MERCOSUR dummy	-	-	-	0.62**	0.79***
				(2.09)	(2.61)
NAFTA dummy	-	-	-	-0.21	0.10
				(-0.67)	(0.26)
CAN dummy	-	-	0.12	0.94***	1.16***
			(0.53)	(4.26)	(3.94)
UE dummy	0.68***	0.40***	0.42***	0.36***	0.28***
	(5.02)	(2.95)	(3.65)	(4.82)	(3.59)
Distance	-1.07***	-1.15***	-1.03***	-1.20***	-1.31***
	(-14.08)	(-16.28)	(-16.03)	(-21.15)	(-23.36)
DP*Distance	**0.14**	**0.27*** **	**0.29*** **	**0.38*** **	**0.43*** **
	(1.32)	**(2.75)**	**(3.35)**	**(5.25)**	**(6.02)**
Language dummy	0.48***	0.48***	0.72***	0.77***	0.80***
	(3.45)	(3.42)	(5.61)	(6.87)	(7.41)
DP*Language dummy	**-0.22**	**-0.46** **	**-0.66*** **	**-0.41** **	**-0.39** **
	(-0.99)	**(-2.04)**	**(-3.07)**	**(-2.37)**	**(-2.31)**
Exporter's ArCo	4.57***	3.27***	4.36***	4.49***	4.87***
	(7.12)	(4.75)	(6.70)	(9.95)	(11.49)
DP*Exporter's ArCo	**0.45**	**-0.03**	**-2.63*** **	**-2.64*** **	**-2.68*** **
	(0.53)	**(-0.03)**	**(-3.04)**	**(-4.49)**	**(-4.81)**
Importer's ArCo	2.99***	3.49***	2.34***	2.25***	1.91***
	(4.42)	(4.81)	(3.70)	(4.39)	(3.88)
DP*Importer's ArCo	**-1.45**	**-1.63***	**-1.88** **	**-1.87*** **	**-1.17***
	(-1.61)	**(-1.73)**	**(-2.22)**	**(-2.99)**	**(-1.91)**
R-squared	0.65	0.67	0.71	0.78	0.80
Number of observations	2440	2408	2926	3334	3347
Akaike Info Criterion	22.91	22.87	23.76	24.49	24.52
Root Mean Squared Error	1.75	1.73	1.67	1.49	1.46
Mean Absolute Error	1.32	1.29	1.24	1.11	1.09
Mean Absolute Percentage Error	20.76	20.41	18.36	15.33	15.11
S&J (1990) goodness of fit	0.91	0.91	0.92	0.93	0.93

Notes: ***, **, * indicate significance at 1%, 5% and 10%, respectively. T-statistics are in brackets. The dependent variable is the natural logarithm of exports in value (current US$). Income, population and distance are also in natural logarithms. The estimation uses White's heteroscedasticity-consistent standard errors. DP is a dummy variable that takes the value one when the exporter country is a high income country and zero otherwise.

Table 3. Determinants of international trade. Augmented non-linear gravity model.

	1980	1985	1990	1995	1999
Exporter's income	0.61***	0.77***	1.00***	0.62***	0.56***
	(12.76)	(14.13)	(8.94)	(6.56)	(5.96)
Importer's income	0.67***	0.87***	0.98***	0.72***	0.70***
	(19.44)	(17.27)	(12.66)	(8.91)	(10.25)
Exporter's population	0.83***	0.89***	0.72***	0.75***	0.83***
	(12.55)	(10.38)	(12.75)	(12.70)	(15.50)
Importer's population	0.87***	1.09***	0.71***	0.93***	1.10***
	(9.81)	(11.66)	(6.35)	(10.88)	(11.65)
Adjacency dummy	0.84***	0.73***	1.60***	1.19***	1.12***
	(9.20)	(7.48)	(9.94)	(6.04)	(6.29)
Exporter island dummy	1.46***	1.71***	0.68***	0.53***	0.59***
	(8.85)	(9.73)	(4.62)	(2.78)	(3.32)
Importer island dummy	1.05***	1.28***	0.68***	0.44***	0.49***
	(8.46)	(9.14)	(3.88)	(2.71)	(3.86)
Exporter landlocked dummy	-0.64***	-0.43**	-0.44**	-0.02	0.17
	(-3.58)	(-2.03)	(-2.35)	(-0.09)	(0.88)
Importer landlocked dummy	1.07***	0.89***	-0.27	0.37**	0.42**
	(4.73)	(3.41)	(-1.01)	(1.98)	(2.07)
Exporter's remoteness	-0.60**	-0.60*	-0.20	-0.24	-0.41
	(-2.05)	(-1.75)	(-0.52)	(-0.67)	(-1.35)
Importer's remoteness	-0.86***	-1.52***	-2.32***	-1.51***	-1.50***
	(-3.20)	(-3.86)	(-5.52)	(-3.08)	(-3.74)
CACM dummy	-	-	4.40***	2.55***	3.43***
			(3.84)	(2.73)	(4.33)
CARICOM dummy	3.66***	3.55***	5.23***	4.03***	4.17***
	(6.33)	(5.34)	(5.48)	(4.03)	(3.94)
MERCOSUR dummy	-	-	-	1.99***	2.15***
				(4.66)	(5.55)
NAFTA dummy	-	-	-	0.40**	0.90***
				(2.21)	(4.95)
CAN dummy	-	-	2.22***	2.01***	1.68***
			(2.70)	(3.18)	(3.00)
UE dummy	1.80***	1.37***	1.06***	0.59***	0.82***
	(6.37)	(6.20)	(3.29)	(3.99)	(5.31)
Distance	0.51***	0.74***	0.13	0.23***	0.18***
	(6.86)	(9.96)	(1.52)	(3.58)	(2.75)
Language dummy	-0.16	-0.35**	-0.48***	-0.11	-0.26**
	(-1.45)	(-2.52)	(-3.09)	(-0.76)	(-2.20)
Exporter's ArCo	6.44***	4.83***	4.45***	3.44***	3.17***
	(6.88)	(5.48)	(5.02)	(5.54)	(6.96)
Importer's ArCo	5.60***	7.74***	7.24***	6.16**	8.45***
	(4.00)	(4.27)	(2.93)	(2.15)	(2.91)
R-squared	0.93	0.95	0.85	0.83	0.87
Number of obs	2999	2997	3447	3689	3692
Number of iterations	45	42	41	39	28
Akaike Info Criterion	28.74	29.22	31.11	32.23	32.34
Root Mean Squared Error	419152.1	533166	1367821	2397003	2529591
Mean Absolute Error	124008.2	132031.3	308903.2	605801.6	660928.4
Mean Absolute Percentage Error	1259.01	812.14	1079.99	3746.58	4054.8
S&J (1990) goodness of fit	0.87	0.89	0.81	0.79	0.82

Notes: ***, **, *, indicate significance at 1%, 5% and 10%, respectively. T-statistics are in brackets. The dependent variable is exports in value (current US$). The estimation uses White's heteroscedasticity-consistent standard errors.

APPENDIX

Table A.1: Variable descriptions and sources of data.

Variable	Description	Source
X_{ij} : Exports from i to j	Nominal value of bilateral exports	Statistics Canada (2001)
Y_i : Exporter's income	Exporter's GDP, PPP (current international $)	World Bank (2001)
Y_j : Importer's income	Importer's GDP, PPP (current international $)	World Bank (2001)
P_i : Exporter's population	Total population in the exporter's market	World Bank (2001)
P_j : Importer's population	Total population in the importer's market	World Bank (2001)
Adj_{ij} : Adjacency dummy	Dummy variable = 1 if the trading partners share a border, 0 otherwise	CIA (2003)
Isl: Island dummy	Dummy variable = 1 if the country is an island, 0 otherwise	CIA (2003)
Land: Landlocked dummy	Dummy variable = 1 if the country is landlocked, 0 otherwise	CIA (2003)
Remoteness	Variable that measures the level of remoteness in the exporter country.	Coca-Castaño et al. (2005) Authors' calculations
CACM dummy	Dummy variable = 1 if the trading partners are members of CACM, 0 otherwise	
CARICOM dummy	Dummy variable = 1 if the trading partners are members of CARICOM, 0 otherwise	
MERCOSUR dummy	Dummy variable = 1 if the trading partners are members of MERCOSUR, 0 otherwise	
NAFTA dummy	Dummy variable = 1 if the trading partners are members of NAFTA, 0 otherwise	
CAN dummy	Dummy variable = 1 if the trading partners are members of CAN, 0 otherwise	
UE dummy	Dummy variable = 1 if the trading partners are members of European Union, 0 otherwise	
$Dist_{ij}$: Distance	Great circle distances between country capitals of trading partners (km)	Great circle distances between cities (2003)
$Lang_{ij}$: Language dummy	Dummy variable = 1 if the trading partners share the same official language, 0 otherwise.	CIA (2003)
$ArCo_i$: Exporter's ArCo	Technological variable	Archibugi and Coco (2002)
$ArCo_j$: Importer's ArCo	Technological variable	Archibugi and Coco (2002)
Inf_i : Exporter's infrastructure	Transport infrastructure variable	CIA (2003), authors' calculations
Inf_j : Importer's infrastructure	Transport infrastructure variable	CIA (2003), authors' calculations

Note: UNDP denotes United Nations Development Programme and CIA denotes Central Intelligence Agency.

Table A.2: Evolution of remoteness in the 65-country sample.

	1980	1985	1990	1995	1999
South Africa	7600.21	7689.27	8192.48	9029.90	9132.98
Algeria	4587.57	4677.42	4846.33	5310.92	5426.24
Argentina	7574.29	7868.63	8724.80	9991.96	10200.52
Australia	9582.12	9521.41	10267.41	11336.27	11356.86
Austria	4471.43	4511.79	4598.19	4931.49	5027.47
Belgium-Luxembourg	4268.07	4332.14	4415.58	4782.51	4887.85
Bolivia	6676.63	6946.53	7738.05	8906.91	9098.98
Brazil	6527.71	6795.66	7520.06	8601.50	8786.50
Bulgaria	4721.65	4750.24	4857.93	5219.87	5307.36
Canada	4009.45	4128.08	4555.27	5279.33	5377.15
Chile	7576.15	7860.71	8737.71	10033.56	10243.21
China	6180.10	5968.73	6142.68	6517.45	6446.15
Colombia	5618.26	5842.82	6525.68	7551.67	7714.07
Costa Rica	5480.04	5679.84	6353.63	7363.05	7508.80
Croatia	4509.03	4554.47	4651.03	5003.47	5101.70
Cyprus	5161.47	5165.69	5303.65	5698.37	5768.09
Denmark	4344.47	4380.82	4455.12	4774.51	4866.06
Dominican Republic	4899.46	5095.32	5679.18	6587.99	6730.05
Ecuador	5923.38	6152.46	6875.60	7949.96	8116.12
Egypt	5305.39	5320.42	5484.77	5914.29	5988.90
Slovak Republic	4486.30	4524.72	4611.21	4944.29	5039.20
Finland	4484.70	4488.54	4568.60	4888.54	4958.83
France	4269.23	4341.80	4436.47	4825.24	4934.65
Germany	4374.46	4415.10	4490.87	4796.98	4893.22
Ghana	5832.06	5968.71	6350.90	7052.17	7179.93
Greece	4868.75	4901.10	5031.31	5423.06	5510.84
Hong Kong, China	7075.61	6878.79	7124.93	7622.99	7573.31
Honduras	5325.31	5510.91	6165.33	7149.17	7285.42
Iceland	4156.21	4212.63	4378.82	4828.29	4917.00
India	6245.28	6111.74	6270.59	6671.46	6649.65
Ireland	4199.35	4272.87	4383.64	4796.94	4902.09
Israel	5323.05	5324.07	5477.02	5889.58	5956.15
Italy	4538.33	4599.46	4720.02	5112.25	5217.76
Jamaica	4977.70	5168.59	5775.22	6706.88	6845.46
Japan	6373.55	6191.92	6460.60	6993.89	6969.01
Kenya	6755.30	6787.22	7120.90	7757.81	7835.27
Rep. Korea	6325.71	6121.66	6334.87	6781.39	6729.95
Mexico	5329.84	5481.10	6120.60	7080.81	7196.71
Mozambique	7700.07	7776.09	8271.60	9100.07	9197.44
Netherlands	4266.77	4326.33	4406.54	4764.26	4867.29
Nicaragua	5408.39	5599.91	6264.88	7262.20	7402.46
Norway	4317.24	4348.03	4431.85	4769.70	4855.07
Nepal	6445.90	6294.06	6466.76	6879.88	6848.48
Panama	5461.20	5669.73	6339.46	7346.25	7497.59
Peru	6528.42	6781.27	7569.53	8728.65	8911.57
Pakistan	5978.88	5858.05	5999.97	6379.28	6366.93
Poland	4496.02	4519.05	4592.48	4905.25	4990.30
Portugal	4445.20	4555.51	4752.08	5259.90	5380.83
Paraguay	7084.46	7369.94	8179.09	9375.63	9574.79
Czech Republic	4414.66	4457.16	4538.14	4860.30	4957.55
El Salvador	5382.91	5565.94	6226.41	7216.75	7351.66

Senegal	5272.80	5438.29	5829.25	6544.12	6684.44
Singapore	8091.13	7939.97	8278.31	8904.39	8884.19
Spain	4419.22	4517.92	4682.87	5153.79	5271.43
Sudan	5930.64	5956.71	6197.87	6722.77	6799.52
Sweden	4398.47	4417.16	4495.83	4818.10	4897.65
Syrian Arab Republic	5283.16	5279.42	5423.80	5824.85	5888.92
Switzerland	4357.88	4423.30	4521.76	4895.59	5003.01
Tanzania	7050.50	7081.44	7444.78	8117.28	8193.98
Trinidad and Tobago	5245.75	5464.34	6070.30	7010.51	7168.18
Turkey	4992.59	4994.62	5111.71	5479.92	5550.04
United Kingdom	4235.16	4305.27	4388.66	4778.21	4885.12
Uruguay	7586.78	7879.03	8729.12	9986.97	10192.38
United States	4001.08	4137.96	4614.19	5390.73	5493.84
Venezuela	5268.30	5485.66	6110.94	7070.76	7227.42

Table A.3. Comparison of non-linear coefficients.

Variable	Table 3[a]	Coe et al. (2002)[b]	Croce et al. (2004)[c]
Exporter's income	0.56***	-	-
Importer's income	0.70***	-	-
Economic mass	-	0.74**	0.68***
Exporter's population	0.83***	-	0.51***
Importer's population	1.10***	-	0.59***
Population	-	0.11	-
Adjacency dummy	1.12***	0.52**	1.12***
Exporter's remoteness	-0.41	-	-
Importer's remoteness	-1.50***	-	-
Remoteness	-	0.46	-
CACM dummy	3.43***	-	1.53
MERCOSUR dummy	2.15***	-	2.08***
NAFTA dummy	0.90***	-	0.28***
CAN dummy	1.68***	-	1.86**
Distance	0.18***	-0.32**	-0.49***
Language dummy	-0.26**	0.03	0.12***
Adjusted R-squared	0.87	0.91	0.97
Number of observations	3692	2613	4017

Notes: ***, **, *, indicate significance at 1%, 5% and 10%, respectively. Coe et al. (2002) and Croce et al. (2004) estimate a cross section for different years. The years nearer to 1999 have been selected to compare.

a. 65-country sample. Year 1999. The dependent variable is exports in value (current US$).

b. 73-country sample. Year 2000. The dependent variable is imports in value (current US$ billions).

c. 64-country sample. Year 1998. The dependent variable is total trade in dollars.

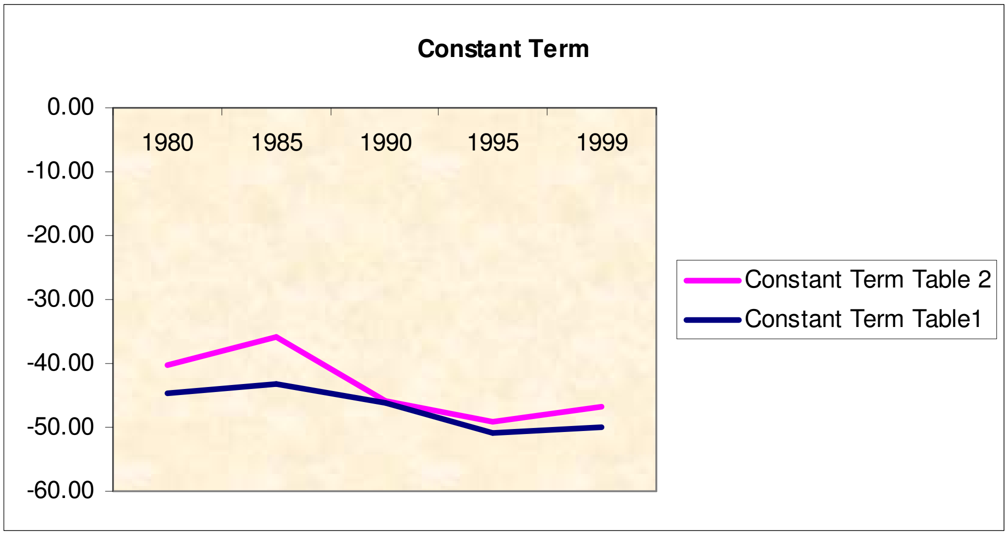
Figure 1. Selected countries.

South Africa
Algeria
Argentina
Australia
Austria
Belgium-Luxembourg
Bolivia
Brazil
Bulgaria
Canada
Chile
China
Colombia
Costa Rica
Croatia
Cyprus
Denmark
Dominican Republic
Ecuador
Arab Rep. Egypt
Slovak Republic
Finland
France
Germany
Ghana
Greece
Hong Kong, China
Honduras
Iceland
India
Ireland
Israel
Italy
Jamaica
Japan
Kenya
Rep. Korea
Mexico
Mozambique
Netherlands
Nicaragua
Norway
Nepal
Panama
Peru
Pakistan
Poland
Portugal
Paraguay
Czech Republic
El Salvador
Senegal
Singapore
Spain
Sudan
Sweden
Syrian Arab Republic
Switzerland
Tanzania
Trinidad and Tobago
Turkey
United Kingdom
Uruguay
United States
Venezuela

Developed countries: South Africa, Argentina, Australia, Austria, Belgium-Luxembourg, Canada, Chile, Costa Rica, Cyprus, Denmark, Slovak Republic, Finland, France, Germany, Greece, Hong Kong, Iceland, Ireland, Israel, Italy Japan, Rep. Korea, Netherlands, Norway, Poland, Portugal, Czech Republic, Singapore, Spain, Sweden, Switzerland, United Kingdom, Uruguay, United States

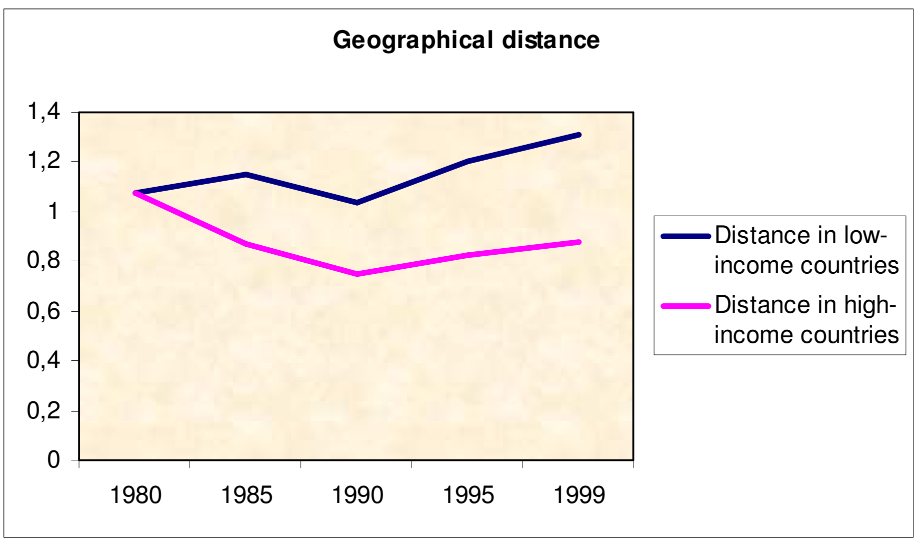
Developing countries: Algeria, Bolivia, Brazil, Bulgaria, China, Colombia, Croatia, Dominican Republic, Ecuador, Arab Rep. Egypt, Ghana, Honduras, India, Jamaica, Kenya, Mexico, Mozambique, Nicaragua, Nepal, Panama, Peru, Pakistan, Paraguay, El Salvador, Senegal, Syrian Arab Republic, Tanzania, Trinidad and Tobago, Turkey, Venezuela.

Figure 2. Evolution of the constant term (from Tables 1 and 2).



Source: Own Elaboration

Figure 3. Evolution of the geographical distance in high and low-income countries in absolute value (from Table 2).



Source: Own Elaboration