University of Sussex
Business, Management & Economics

Economics Department Working Paper Series

No. 12-2010

Is the Border Effect an Artefact of Geographic Aggregation?

Carlos Llano (Universidad Autónoma de Madrid)
Asier Minondo (Universidad de Deusto-ESTE)
Francisco Requena (Universidad de Valencia)

Abstract: The existence of a large border effect is considered as one of the main puzzles of international macroeconomics. We show that the border effect is, to a large extent, an artefact of geographic concentration. In order to do so we combine international flows with intranational flows data characterised by a high geographic grid. At this fine grid, intra-national flows are highly localised and dropping sharply with distance. The use of a small geographical unit of reference to measure intra-national bilateral trade flows allows to estimating correctly the negative impact of distance on shipments. When we use sector disaggregated export flows of 50 Spanish provinces in years 2000 and 2005 split into interprovincial and inter-national flows, we find that the border effect is reduced substantially and even becomes statistically not different from zero in some estimations.

JEL Classification: F14, F15
Key Words: border effect, distance, interregional trade, international trade, Spanish provinces

Acknowledgements: We are grateful to Robert Feenstra, Dennis Novy, Deborah Swenson, James Walker and participants of ETSG Conference in Rome (September 2009), IAES Conference in Boston (October 2009), North American Meeting of the Regional Science Association International in San Francisco (November 2009), SAE Conference in Valencia (December 2009) and Rocky Mountain Empirical Trade Conference in Banff (May 2010). Carlos Llano acknowledges financial support from the Education Department of the Regional Government of Madrid (project TransporTrade S2007/HUM/497 www.uam.es/transportrade). Asier Minondo acknowledges financial support from the Department of Education, Universities and Research of the Basque Government. Francisco Requena acknowledges financial support from the Spanish Ministry of Science and Innovation (project number ECO 2008-04059/ECON) and Generalitat Valenciana, project Prometeo/2009/098.

1. Introduction

Since McCallum's (1995) seminal paper, a large body of literature has tried to explain why countries trade much more with themselves than with other partners. The notoriety of the border effect is such that it has been argued to be “one of the six major puzzles in international macroeconomics” (Obstfeld and Rogoff, 2000). Researchers have followed four different routes to solve the boprder puzzle. First, some authors (Rauch, 2001; Combes et al., 2005; Schulze and Wolf, 2009; Nitsch and Wolf, 2009) argue that large border effects should not constitute a puzzle, because information barriers increase gently the costs of doing business not only for partners located in different countries but even for patterns located in different regions within a country. Second, other authors point out that even when trade costs were low, a large elasticity of substitution between domestic and foreign goods may lead to large home-bias effects (Evans, 2003). Chaney (2008) demonstrates that the presence of fixed costs of exporting and productivity heterogeneity across firms also magnify the aggregated trade elasticity with respect to trade barriers. Rossi-Hansberg (2005) and Yi (2009) show how endogenous firm location and vertical specialisation in the presence of intermediate goods may help to magnify initial small trade frictions along national borders. A third group of researchers argue that large border effects arise due to a misspecification of the gravity model used in the econometric analyses (Anderson and van Wincoop, 2003) or to identification problems (Heinemeyer et al, 2008).

Finally, a fourth group of authors have indicated that the mismeasurement of physical distance may explain the large border effects. This concern arose in studies that lacked data on flows across sub-national units and had to input intra-national trade as the difference between production and exports (Wei, 1996; Nitsch, 2000; Head and Mayer, 2000; Helliwell and Verdier, 2001). In these studies the distance travelled by goods in the intra-national flow plays a crucial role. If the internal distance is not estimated accurately, the estimated border effect will not be reliable. In particular, if intra-national distance is overestimated, a larger border effect will be needed to explain the preference for national goods with respect to foreign goods. Head and Mayer (2002) show that most studies that used imputed intra-national flows overestimate intra-national distance, leading to illusory border effects.

The distance mismeasurement problem does not arise, in principle, in those studies, such as McCallum (1995), that combine intra-national flows across sub-national units and across those sub-national units and foreign sub-national units (or countries), as long as the same procedure is followed to calculate intra-national and inter-national distances. However, these studies may also face limitations if they use geographically large sub-national units, such as US states or Canadian provinces. In a recent work, Hillberry and Hummels (2008) shows that the relationship between shipments and distance is highly non-linear: at the beginning, there is a sharp reduction in value with distance; however, once a distance-threshold is achieved the negative effect vanishes. If most of intra-national trade happens at short distances, the choice of the sub-national unit will play a crucial role in the border effect analysis. 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. Hence, econometric analyses may underestimate the negative contribution of distance on trade, and a larger border effect will be needed to explain the preference for national goods with respect to international goods.

Hillberry and Hummels (2008) analyse whether aggregation may explain why previous studies, such as Wolf (2000) and Hillberry and Hummels (2003), find that US states trade more with themselves than with other US states. They show that the intra-national home bias vanishes once geographically finely disaggregated data, which picks the non-linear relationship between distance and shipments value, are used. Our paper extends Hillberry and Hummels´ findings to the international trade level and analyse to what extent large border (frontier) effects can also be an artefact of geographical aggregation. For that purpose we use a unique database that reports the value of Spanish small geographical units (provinces) trade flows to other small geographical units and to other countries in years 2000 and 2005. Previous research by Gil et al. (2005) use trade flows between Spanish large geographical units (regions) to the rest of Spain and to other countries over the period 1995-1998. They found that Spanish regions traded 21.8 times more with the rest of Spain than they did with other foreign countries, a result very similar to that found by McCallum for Canada. We show that the Spanish border effect is reduced substantially, and even disappears, once we allow data to show the high geographic concentration of intra-national trade flows. These results are in line with recent research, such as Gorodnichenko and Tesar (2009), which using price rather than quantities has also put into question the excessively large border effects found by previous studies.

The rest of the paper is organised as follows. Section 2 analyses whether there is a non-linear relationship between distance and the value of trade flows in Spain. Section 3 presents the empirical model and Section 4 the main results of the econometric analyses. The final section summarises the main conclusions of our paper.

2. The geographic concentration of trade flows

Aggregation will lead to an overestimation of the border effect as long as there is a non-linear relationship between distance and the value of trade flows. To document this feature in the case of Spain, we use a new database, *C-Intereg*, which combines fine grid intra-national trade flows data with international exports data (Llano, 2004).[1] It offers the value of total trade flows from Spanish provinces (Eurostat NUTS III) to other Spanish provinces and to 17 European countries, disaggregated by 13 manufacturing industries.[2] Distance data is obtained from the Spanish Permanent Survey on Commodity Transport by Road (SPSCTR), both for internal and external trade flows. Our distance measure is the actual average distance travelled by heavy trucks in their deliveries of commodities for the inter-provincial and international flows. This measure of distance has the virtuosity of capturing the distance travelled by trucks between actual origins and destinations.[3] Spain is divided into 50 provinces, a relatively small

---

[1] The C-intereg database (see www.c-intereg.es) combines the most accurate data on Spanish transport flows of goods by transport modes (road, rail, ship, plane, pipe and electric network) with additional information used to estimate specific export price vectors, province of origin, transport mode and type of product . The methodology also includes a process for correcting problems in the original transport flows database that arise from multi-modal transport flows and international transit flows hidden in the interregional flows. After this first screening process, an initial estimate of interregional trade flows in tons and in current euro is obtained, based on a combination of transport magnitudes and price information. Finally, a process of harmonization is applied to produce flows magnitudes in tons and euros that are consistent with total output magnitudes from the Spanish Industrial Survey, the National Accounts and the international trade figures from Customs. This final process of harmonization allows offering international trade flows with origin (destination) in any of the Spanish provinces and destination (origin) in any country of the world, compatible with the internal trade estimates and the official output figures at the provincial and the sector level. This procedure brings the trade flows to a position close to the paradigm of trade in continuous space.

[2] The European countries are Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia and Sweden. The industries are Food, beverages and tobacco; Textiles; Leather and footwear; Wood; Paper, printing and publishing; Chemicals; Rubber and plastics; Other non-metallic minerals; Basic and fabricated metals; Machinery; Electrical and Optical Equipment; Transport equipment; Other manufacturing.

[3] The estimate of the actual distance travelled by the trucks for the intra provincial and interprovincial flows is obtained straightforward from the survey. The distance for the international flows between each of the Spanish provinces and the European countries implies a more complex procedure: the SPSCTR also captures bilateral flows between the Spanish provinces and a large list of countries (with specification of the distance travelled but not of the international region of destination/origin). Although in theory, every international flow carried out by Spanish trucks is covered by the survey, the number of observations for every year may be not representative enough for some bilateral flows. In order to increase the quality of our distance measure, and taking into account the stable nature of this variable, we increase the number of observations by pooling together all the international shipments for the period 1995-2006. Thus, a database of about 200.000 observations regarding the average actual distance for the international deliveries was obtained for each dyad of Spanish provinces and European countries in that period.

administrative unit whose average area is similar to an US 3-digit zip code.[4] Due to the characteristic of our distance measure, we exclude from the sample those intra-national flows where a Spanish island-province is a partner (Balearic Islands, Las Palmas and Tenerife).

We use a kernel regression estimator to provide a nonparametric estimate of the relationship between distance and the value of Spanish intra-national and international trade flows for year 2005.[5] As shown in Figure 1, there is a negative relationship between the value of trade flows and distance. However, this negative relationship does not follow a clean pattern. In a first segment, 25 km. to 900 km., the value of shipments drops almost four times, from 65000 thousand euros to 17000 thousand euros. This reduction is especially sharp at short distances, when distance increases from 25 km. to 250 km. In a second segment, 900 km. to 1200 km., paradoxically, we observe an increase in the value of trade with distance. In a third segment, 1200 km. to 2700 km. there is again a negative relationship between the value of trade flows and distance; finally, the relationship becomes flat from 2700 onwards.

Next we examine separately the relationship between the value trade flows and distance for the intra-national flows and inter-national flows, separately, in year 2005. As shown in Figure 2, we now observe a clear non-linear relationship between the value of trade flows and distance for intra-national flows. The value of intra-national trade flows drop from 120000 thousand euros to 28000 thousand euros when distance increases from 25 km. to 200 km. From this distance onwards the relationship is slightly negative and even flat for some distance ranges.

To analyse the relationship between distance and international trade flows, we divide exports by the product of production at origin and demand at destination (adjusted trade), since international partners are very different in their demand capacity. Figure 3 shows that the non-linearity between the value of adjusted trade flows and distance is much sharper for international shipments. The value of trade flows drops more than 80 times when distance increases from 200 km. to 600 km.; from 600 km. onwards distance does not seem to affect the value of shipments.

[4] The average area of a US 3-digit zip code is 13763 square kilometres and the average area of a Spanish province is 10116 square kilometres (excluding Ceuta and Melilla, two autonomous city-provinces in Africa).
[5] We use the Gaussian kernel estimator in STATA, calculated on n=100 points, and allowing the estimator to calculate the optimal bandwidth.

Finally, Figure 4 presents the relationship between adjusted trade and distance for intra-national flows when we use highly disaggregated geographical units as origin and destination partners (provinces: PROV-PROV line), and when we use a geographically larger destination partner (Rest of Spain: PROV-ROS line).[6] Again the relationship between adjusted trade and distance is highly non-linear relationship when province is the destination partner (PROV-PROV line); on the contrary, when Rest of Spain is used as destination partner the relationship is flat for the whole distance-range (PROV-ROS line).

The comparison of the PROV-PROV line and the PROV-ROS line helps to understand how the use of highly aggregated geographical sub-units may lead to an overestimation of the border effect. The use of aggregated units does not allow capturing the highly non-linear relationship between distance and the value of intra-national shipments, leading to an underestimation of the negative impact of distance on trade at short distances.[7] In this paper we want to examine to what extent such underestimation may lead, in turn, to create an "artificial" border effect. In particular, if distance (and other controls for trade impediments) does not explain accurately interregional trade within a country, then the Spanish frontier will be capturing it, exaggerating the “true” frontier effect. The next sections examine this possibility.

3. Empirical model

We estimate the border effect based on the structural gravity equation developed by Anderson and van Wincoop (2003).[8] The estimating equation takes the following form:

[6] Rest of Spain is the geographical aggregation for the destination sub-national unit used by Gil et al. (2005) in their analysis of the Spanish border effect. As the origin sub-national unit (exporter), they use regions (NUTS - 2). Spanish regions, also called autonomous communities, are aggregations of Spanish provinces. There are 17 regions in Spain, of which seven have only one province. The results are very similar if we use this aggregation for the origin sub-national unit.

[7] In Figure 3 we observe as well that the value of shipments drops sharply with distance for international flows. In this case the use of geographically aggregated national units would underestimate the border effect as the distance travelled by goods in inter-national shipments would be overestimated. In any case, this problem seems to be of a lower magnitude, because differences in travelled distance would only be relevant for countries that are very close to Spain (France and Portugal).

[8] Helpman, Melitz and Rubinstein (2008) develop a structural gravity model that incorporates heterogeneous firms. They argue that previous studies' estimates are biased because they do not take into account the large number of zeros in bilateral trade. As explained in the text, our estimation technique controls for this problem.

$$\ln X_{ijkt} = \beta_0 + \beta_1 Border + \beta_2 \ln Y_{ikt} + \beta_3 \ln Y_{jt} + \beta_4 \ln dist_{ij} + \beta_5 Contig\_Int\_Partner_{ij} + \beta_6 Contig\_Nat\_Partner_{ij} + P_{i,k} + P_{j,k} \quad (1)$$

where X_{ijkt} are exports from province *i* to a trading partner *j* of industry *k* in year *t. Border* is a dummy variable that takes the value of 1 if the shipment is intra-national and zero if the shipment is international; Y_{ikt} is production of industry *k* in province *i* in year *t* (source: INE); Y_{jt} denotes trading partner's GDP *j* in year *t* (source: WDI-online). A dummy variable *Contig_Int_Partner* that takes values of one if the Spanish exporting province is adjacent to an importing international trading partner (a foreign country), zero otherwise, and a dummy variable *Contig_Nat_Partner* that takes values of one if the Spanish exporting province is adjacent to an importing national partner (a province), zero otherwise.[9] We expect these variables to be positive if our sample does not capture the whole non-linear relationship between trade and distance that happens at short distances. To control for multilateral resistances, we include industry-province and industry-country fixed effects (P_i^k, P_j^k).

In order to capture the non-linear relationship between distance and the value of shipments we also estimate the following equation:

$$\ln X_{ijkt} = \beta_0 + \beta_1 Border + \beta_2 \ln Y_{ikt} + \beta_3 \ln Y_{jt} + \beta_5 dist_{ij} + \beta_6 dist^2{}_{ij} \; \beta_6 Contig\_Int\_Partner_{ij} + \beta_7 Contig\_Nat\_Partner_{ij} + P_{i,k} + P_{j,k} \quad (2)$$

where we have introduced a coefficient for distance and another for the square of distance. As explained before, distance is the actual average distance travelled by heavy trucks in their deliveries of commodities for the inter-provincial and international flows. It is important to stress that our distance measure helps to address the aggregation problem both for international and intra-national shipments.

Finally, due to the existence of zeros in our database, as in Eichengreen and Irwin (1998) and Chen (2004), we express the dependent variable as (1+Ln X_{ijkt}) and estimate the model by a Tobit procedure.[10] The equations (1) and (2) are estimated using data from years 2000 and 2005.

[9] When a province sells to the same province the dummy variable takes the value of 1.

[10] We use the McDonald and Moffitt (1980) procedure to recover the elasticity at sample means.

The definition of a Spanish trading partner plays a key role in our exercise. While the number of international trading partners will be constant and equal to 17 countries in all the estimations, the number and economic size of the Spanish trading partners will change. In the first sample we aggregate all Spanish trading partners in one, the rest of Spain, so each exporting province has 18 trading partners: Rest of Spain plus 17 European countries. We call this sample "SPAIN". Our second sample considers the smallest geographic unit that we have as Spanish trading partner, 47 provinces. In this sample, that we call "PROVINCE", an exporting province has 64 trading partners (47 provinces and 17 European countries). Notice that the value of the variable *distance* will change when we choose different definitions of Spanish trading partner. In particular, the value of distance from one province to Rest of Spain is calculated as the simple average of the distance travelled by trucks from that province to other Spanish provinces. We expect a narrow definition of Spanish trading partners to control adequately for the fact that intra-national shipments are highly localised. Trading with bigger partners (i.e. with the rest of Spain as a whole instead of with each province) will tend to upward bias the border effect since the border effect will be capturing part of the non-linear relationship between distance and shipments.

4. Results of the econometric analyses

Table 1 reports the results of estimating Eq (1) and (2) using the two samples based on the definition of Spanish trading partner. The first two columns use the sample SPAIN and the second two columns the sample PROVINCES. When we aggregate intra-national shipments as the rest of Spain, the coefficient associated to the border effect is around 3.4; that is, a Spanish province trades around 30 times more with the rest of Spain than with a European country. This border estimate is slightly larger than the one obtained by Gil et al. (2005). The discrepancy can be explained by the fact that we use different periods, countries and estimation techniques. In the first estimation, where we are assuming a linear (log-log) relationship between trade and distance, the distance coefficient is negative and statistically significant. In the second estimation, we assume a non-linear relationship between trade and distance; as expected, the distance coefficient has a strong negative value and the square of distance coefficient a positive sign, denoting that trade falls sharply at short distances and then becomes flat. However, both coefficients are statistically not significant. It seems that the sample SPAIN is capturing the large non-linearity between trade and distance for international shipments (Figure 3); however, as the relationship for intra-national flows is flat (Figure 4; PROV-ROS line), the overall non-linearity is not very robust. With respect to other independent variables, in both specifications the coefficients for economic size of trading partners (exporter's production and importer's GDP) are positive. Finally, the dummy variable on contiguous countries is statistically not significant.

Next, we present the results of estimating Eq. (1) and (2) using the sample PROVINCES, that is, other provinces as trading partners. We observe that the border effect is statistically not different from zero neither in Column (3) nor in Column (4). Hence, according to our estimates, Spanish provinces trade as much with other Spanish provinces as with other European countries, once we control for other trade impediments. As argued in this paper, the disappearance of the border effect is explained by the use of a more geographically disaggregated sample (SPAIN), which allows capturing the high concentration of shipments at low distances. As shown in Column (4), now both the distance and square of distance coefficients are statistically significant, confirming the high non-linearity between trade and distance. In addition the coefficient on contiguity between provinces is positive and statistically significant. As exports to contiguous province encompass the shortest distance shipments, the positive coefficient may point out that our sample still does not capture the whole non-linearity between trade and distance. Surprisingly, the coefficient on contiguous countries is negative and statistically significant. Finally, as in the previous estimations, the origin province output and the destination partner demand coefficients are positive.

To sum up, we can conclude that the border effect disappears when the definition of Spanish trading partner changes from a highly aggregated geographical unit (rest of Spain as a whole) to a much smaller unit (provinces). Hence, our results indicate that the border effect is, to a large extent, an artifact of the geographic aggregation of intra-national trade flows.

Table 2 reports the industry-specific border coefficients for 13 manufacturing sectors. In both samples, SPAIN (column 1) or PROVINCES (column 2), there are large differences in the border coefficients across industries. If we focus on the first sample (SPAIN - Column 1), we can observe certain heterogeneity across industries: five sectors exhibit very large border effects (metal products, transport equipment, mechanical machinery, plastic & rubber and non-metallic mineral products) and two sectors do not have border effects (textile & apparel and leather & shoes). When we use the PROVINCE sample, the border coefficients tend to become smaller, and in many cases even negative, as we use a smaller geographic unit to define Spanish trading partners.

Robustness analyses

As mentioned above, the border coefficient we obtain in Table 1 (Columns 1 and 2) is larger than the one obtained by Gil et al. (2005) using the same sub-national aggregated destination partner: the Rest of Spain. It is important to analyse where those differences come from. In order to do so, we reproduce as close as possible the exercise performed by Gil et al. (2005) using our database. These authors estimate the Spanish border effect using a random effects model and data on aggregate shipments from Spanish regions to the rest of Spain and to 27 OECD countries in the period 1995-1998. Their distance measure is based on geodesic information. We can match all the characteristics of their sample, except for the number of countries, which is reduced to 25 OECD countries in our sample. Table 3 - Column (1) reproduces the results obtained by Gil et al. (2005). Their border effect for Spain is 3.08. Column (2) reproduces Gil et al.'s exercise using our database. We can see that our border effect, 2.75, is now smaller than the one obtained by Gil et al. (2005). Hence, we can conclude that the differences in the Spanish border effect between previous estimates and those reported in Table 1 - Columns (1) and (2), are explained by differences in the sample and the estimation technique.

Second, we analyse the robustness of our results when using alternative estimation methods. As an alternative method for the treatment of zeros in our database, following Silva and Tenreyro (2007), we re-estimate the equation using a Poisson model. In addition to that, we also estimate the equation using OLS. As shown in Table 4, in both models we observe that the border effect experiences a very large reduction when we shift from an aggregated intra-national partner to a more geographically disaggregated intra-national partner. When we use a non-linear functional form for distance, the border effect drops from 3.643 to 1.566 in the Poisson model and from 3.485 to 0.808 in the OLS model. This result is consistent with our previous conclusions. However, the estimated border effects are now statistically significant. According to the Poisson model, even when we control for aggregation, Spanish provinces still trade five times more with another Spanish province than with a European country (exp 1.566); according to the OLS model a Spanish province trades two times more with another Spanish province than with another European country (exp. 0.808). Notice that, when we use the PROVINCES sample and using the Poisson estimation or the OLS estimation, distance and the square of distance have the opposite signs found in Table 1. In the Poisson model (column 4) the square of distance is negative and in the OLS model (column 7) distance is positive and the square of distance negative.

Third, we test whether our results are robust to changes in the exporting geographic unit. In the new sample, instead of provinces, we select regions as the exporting geographic unit; in this sample, the number of geographic units is reduced from 47 to 15. As shown in Table 5, when we use the sample SPAIN (columns 1-2), the border coefficient is 3.100 in linear distance specification and 2.785 in the non-linear distance specification, suggesting that regions trade around 16-22 times more with the rest of Spain than with another country. These border effects are lower than the ones reported in Table 1. When we use the sample REGIONS (columns 3-4), as expected, we find a large reduction in the border effect: it drops from 3.100 to 0.549 in the linear specification, and from 2.785 to 1.230 in the non-linear specification. Therefore, the border effect also falls when we use a more aggregated geographical unit. However, such a reduction in the border effect is not as large compared to the sample that uses provinces as geographic unit of reference.

4. Conclusions

During the last decades a large number of studies have tried to explain why countries trade more with themselves than with other countries. We show that the border effect in Spain is, to a large extent, an artefact of geographical aggregation. Based on a geographic fine grid data, we find that intra-national trade flows are highly localised, with value dropping sharply with distance. If intra-national trade flows are aggregated this localised pattern is lost, which leads to an overestimation of the distance travelled by goods sold domestically and to the creation of an artificial border effect. Our econometric estimations confirm this fact. While a sizable border effect exists for Spain when intra-national partners are aggregated, the border effect decreases substantially, and even disappears in some estimations, once when we allow data to show the localised nature of intra-national trade flows.

References

Anderson, J.E. and Van Wincoop, E. (2003). “Gravity with Gravitas: A Solution to the Border Puzzle”, *American Economic Review,* **93**, 1, 170-192.

Chaney, T. (2008). “Distorted Gravity: The Intensive and Extensive Margins of International Trade”, *American Economic Review*, 98(4), 1707-1721.

Chen, N. (2004). “Intra-National Versus International Trade in the European Union: Why Do National Borders Matter?”, *Journal of International Economics*, **63**, 1, 93-118.

Combes, P. P., Lafourcade, M. and Mayer, T. (2005) "The trade-creating effects of business and social networks: evidence from France", *Journal of International Economics*, 66: 1–29.

Eichengreen, B. and Irwin, D. (1998). “The Role of History in Bilateral Trade Flows”, in Frankel, J. (ed.), *The Regionalizatrion of the World Economy,* University of Chicago Press, Chicago.

Evans, C.L. (2003). "The Economic Significance of National Border Effects", *American Economic Review*, **93**, 4, 1291-1312.

Gil-Pareja, S., Llorca-Vivero, R., Martínez Serrano J.A. and Oliver-Alonso, J. (2005). "The Border Effect in Spain", *The World Economy*, **28**, 11, 1617-1631.

Gorodnichenko, Y. and Tesar, L. (2009). "Border effect or country effect? Seattle may not be sor far from Vancouver after all", *American Economic Journal: Macroeconomics*, **1**, 1, 219-241.

Head, K. and Mayer, T. (2000). "Non-Europe: the causes and magnitudes of market fragmentation in the EU", *Weltwirtschaftliches Archiv*, **136**, 2, 285-314.

Head, K. and Mayer, T. (2002). "Illusory Border Effects: Distance mismeasurement inflates estimates of home bias in trade", *CEPII Working Paper 2002-01*.

Heinemeyer, H., Chulze, M.-H. and Wolf, N. (2008), “Endogenous Borders? The Effects of New Borders on Trade in Central Europe 1885-1933”, CESifo Working Paper No. 2246

Helliwell, J.F. and Verdier, G. (2001). "Measuring internal trade distances: a new method applied to estimate provincial border effects in Canada", *Canadian Journal of Economics*, **34**, 4, 1024-1041.

Helpman, E., Melitz, M and Rubinstein, Y. (2008). "Estimating Trade Flows: Trading Partners and Trading Volumes", *Quarterly Journal of Economics*, **123**, 2, 441-487.

Hillberry, R. and Hummels, D. (2003). “Intra-National Home Bias: Some Explanations”, *The Review of Economics and Statistics*, **85**, 4, 1089-1092

Hillberry, R. and Hummels, D. (2008). “Trade responses to geographic frictions: A decomposition using micro-data", *European Economic Review*, **52**, 3, 527-550.

Llano, C. (2004). “The interregional trade in the context of a multirregional input-output model for Spain”, *Estudios de Economía Aplicada*, **22**, 3, 539-576.

McCallum, J. (1995). “National Borders Matter: Canada-U.S. Regional Trade Patterns”, *American Economic Review*, **85**, 3, 615-623.

McDonald, J.F. and Moffitt, R.A. (1980). “The Uses of Tobit Analysis”, *The Review of Economics and Statistics*, **62**, 2, 318-321.

Nitsch, V. (2000). "National borders and international trade: evidence from the European Union", *Canadian Journal of Economics*, **33**, 4, 1091-1105.

Nitsch, V. and Wolf, N. (2009). "Tear down this wall: on the persistence of borders in trade", *CEPR Working Paper 7545*.

Obstfeld, M. and Rogoff, K. (2000). “The Six Major Puzzles in International Macroeconomics. Is there a Common Cause?”, *NBER Working Paper 7777*, National Bureau of Economic Research, Cambridge, MA.

Rauch, J.E. (2001). “Business and Social Networks in International Trade”, *Journal of Economic Literature*, **39**, 4, 1177-1203.

Rossi-Hansberg, E. (2005). “A Spatial Theory of Trade”. *American Economic Review*, 95(5): 1464–1491.

Silva, J. and Tenreyro, S. (2006). "The log of gravity", *Review of Economics and Statistics*, **88**, 4, 641-658.

Shultze, M.S. and N. Wolf (2008), “On the origins of border effects: insights from the Habsburg Empire”, *Journal of Economic Geography*, 1-20.

Wei, S. (1996). "Intra-National Versus International Trade: How Stubborn are Nations in Global Integration?", *NBER Working Paper 5531*.

Wolf, H.C. (2000). "Intranational home bias in trade", *Review of Economics and Statistics*, **82**, 4, 555-563.

Yi, Kei-Mu (2009). "Can Multi-Stage Production Explain the Border Effect", *American Economic Reveiw*, forthcoming.

Figure 1. Kernel regression: value of intra-national and inter-national trade flows on distance, year 2005

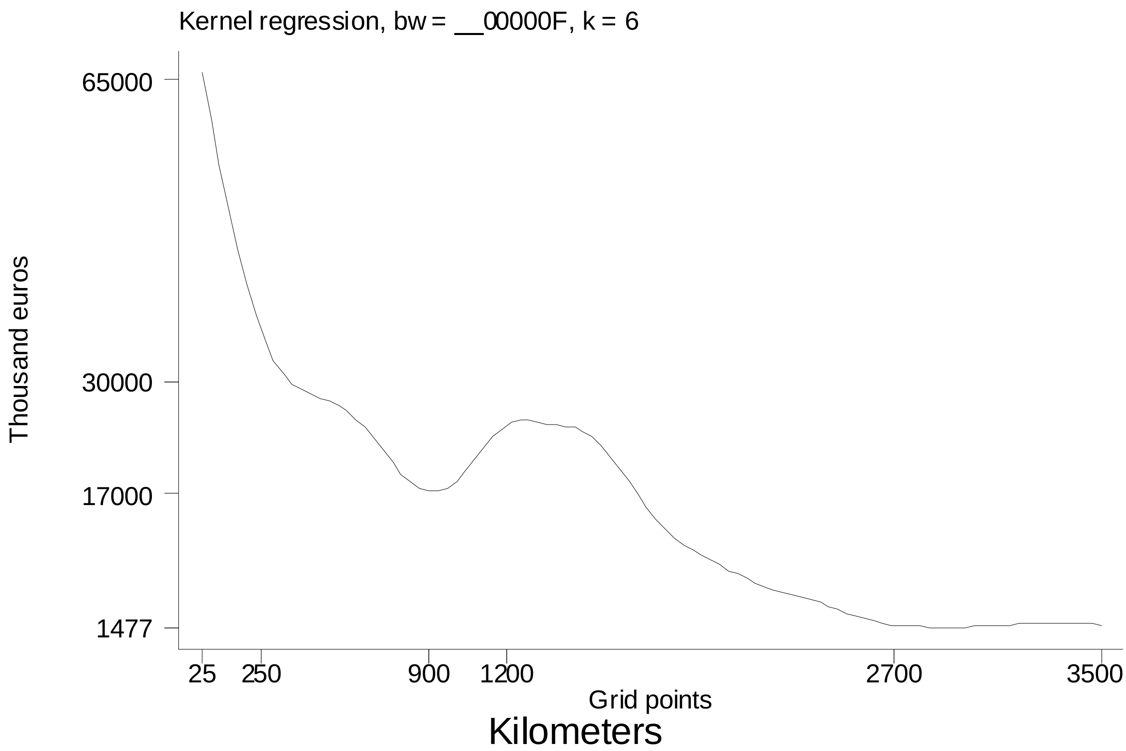

Figure 2. Kernel regression: value of intra-national trade flows on distance, year 2005

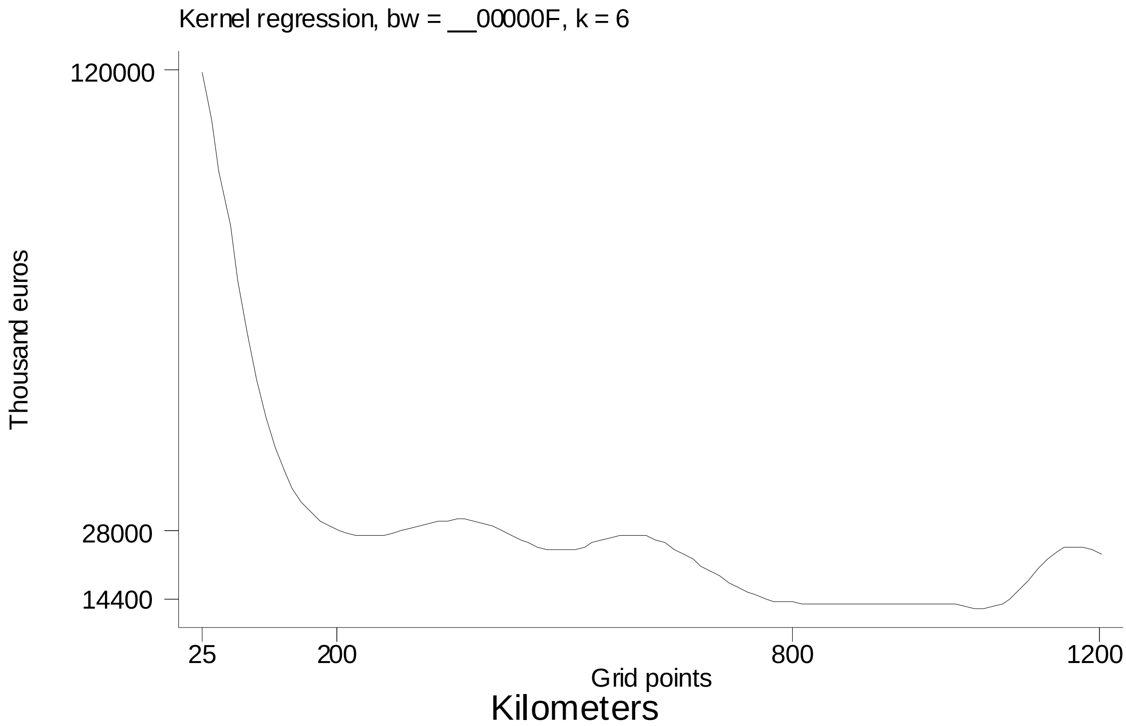

Figure 3. Kernel regression: value of adjusted international trade flows on distance, year 2005

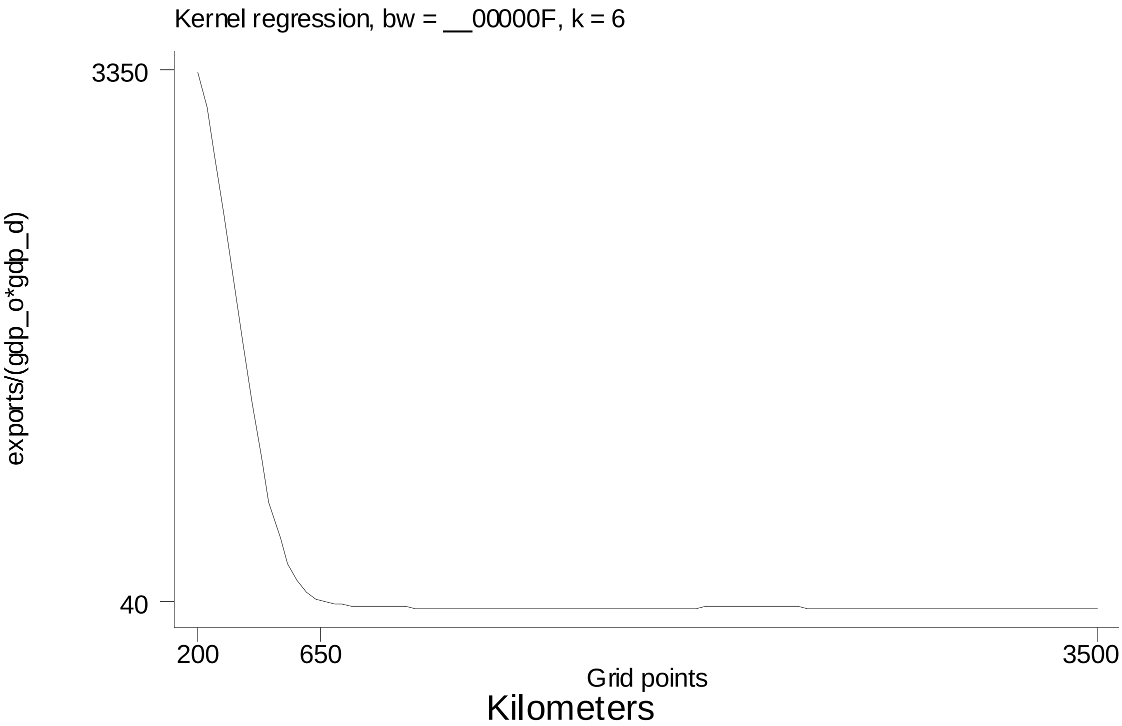

Figure 4. Kernel regression: value of adjusted intra-national trade flows on distance, year 2005

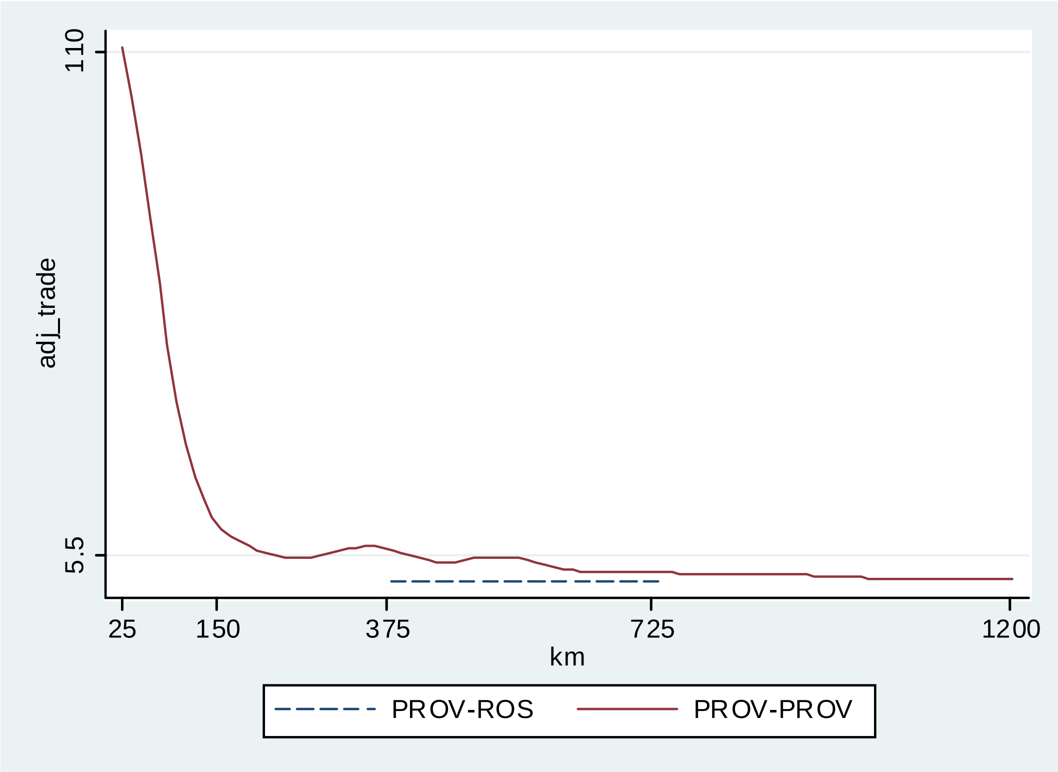

Table 1. Main results

	Sample SPAIN				Sample PROVINCES			
	(1)		(2)		(3)		(4)	
Origin								
Provinces	47		47		47		47	
Destination								
Provinces	-		-		47		47	
Rest of Spain	1		1		-		-	
Countries	17		17		17		17	
Sectors	13		13		13		13	
Years	2		2		2		2	
N	21996		21996		78208		78208	
Border	3.514	***	3.521	***	0.098		0.497	**
	(11.52)		(11.58)		(0.58)		(2.53)	
Gross output exp.	0.280	***	0.280	***	0.147	***	0.147	***
	(3.94)		(3.94)		(3.79)		(3.78)	
GDP importer	1.170	***	1.169	***	0.783	***	0.802	***
	(21.02)		(19.36)		(24.16)		(22.66)	
Distance	-1.429	***	-1.291		-1.130	***	-2.404	***
	(-9.10)		(-0.67)		(-14.34)		(-5.37)	
Distance^2			-0.010				0.109	***
			(-0.07)				(3.02)	
Cont. Int. Part	0.091		0.108		-0.381	***	-0.268	**
	(0.30)		(0.28)		(-2.79)		(-1.98)	
Cont. Nat. Part					0.664	***	0.482	***
					(6.51)		(5.09)	
Pseudo R2	0.301		0.301		0.191		0.192	

Note: Tobit estimations, sample mean elasticities. All specifications include industry-specific exporter, industry-specific importer and year dummies. Province-clustered robust t-statistics in parenthesis with *** denoting significance at the 1-percent level, ** significance at the 5-percent level and * significance at 10-percent level.

Table 2. Industry-specific border effect.

	Sample SPAIN (1)		Sample PROVINCES (2)	
Origin				
Provinces	47		47	
Destination				
Provinces	-		47	
Rest of Spain	1		-	
Countries	17		17	
Sectors	13		13	
Years	2		2	
Observations	21996		78208	
Border Food & drinks	3.335	***	0.600	***
	(15.11)		(2.67)	
Border Textile & apparel	-0.037		-1.478	***
	(-0.09)		(-4.90)	
Border Leather & shoes	-0.199		-1.123	**
	(-0.30)		(-2.38)	
Border Wood products	2.354	***	-1.730	***
	(6.11)		(-6.80)	
Border Paper & print	3.507	***	-0.474	
	(9.71)		(-1.31)	
Border Chemical products	2.203	***	-0.103	
	(6.04)		(-0.34)	
Border Plastic & rubber	5.011	***	-0.458	
	(12.09)		(-1.26)	
Border Nonmetal mineral	4.437	***	0.212	
	(11.38)		(0.67)	
Border Metal products	6.101	***	1.306	***
	(14.61)		(3.53)	
Border Mechanical machinery	5.211	***	0.784	***
	(13.21)		(2.66)	
Border Electric & electronic goods	3.470	***	-0.303	
	(8.01)		(-1.02)	
Border Transport equipment	6.223	***	0.232	
	(13.90)		(0.66)	
Border Other manufactures	3.271	***	-1.325	***
	(7.46)		(-4.51)	
Gross Industry Output Exporter	0.209	***	0.119	***
	(4.41)		(3.40)	
GDP Importer	1.174	***	0.780	***
	(19.25)		(23.58)	
Distance	-1.214	***	-2.454	***
	(-0.63)		(-5.60)	
Distance ^ 2	-0.015		0.112	***
	(-0.11)		(3.14)	
Contig. Int. Partner	0.117		-0.291	**
	(0.31)		(-2.17)	
Contig. Nat. Partner			0.443	***
			(5.10)	
Pseudo R2	0.304		0.193	

Note: Tobit estimations, sample mean elasticities. All specifications include industry-specific exporter, industry-specific importer and year dummies. t-statistics in parenthesis with *** denoting significance at the 1-percent level, ** significance at the 5-percent level and * significance at 10-percent level.

Table 3. Robustness analysis I. Comparison with Gil et al. (2005)

	Gil et al. (2005) estimation	Estimation with our database
Border	3.08 (10.27)***	2.75 (8.09)***
GDP exporter	1.08 (36.71)***	1.23 (17.67)***
GDP importer	1.08 (36.80)***	1.07 (24.62)***
Distance	-1.28 (-21.69)***	-1.53 (-20.42)***
Time period	1995-1998	2000 y 2005
Estimation technique	Random effects model	Random effects model
Exporter	Spanish region	Spanish region
Importer	Rest of Spain + 27 OECD countries	Rest of Spain + 25 OECD countries
Data	Aggregated trade	Aggregated trade

Note: Gil el al. (2005) coefficients correspond to their Table 1-Column 1 estimation. All estimations include year-specific dummies. t-statistics in parenthesis with *** denoting significance at the 1-percent level.

Table 4. Robustness analysis II. Estimation with Poisson and OLS models

	POISSON				OLS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Spain	Spain	Provinces	Provinces	Spain	Spain	Provinces	Provinces
Origin								
Provinces	47	47	47	47	47	47	47	47
Destination								
Provinces	-	-	47	47	-	-	47	47
Rest of Spain	1	1	-	-	1	1	-	-
Countries	17	17	17	17	17	17	17	17
Sectors	13	13	13	13	13	13	13	13
Years	2	2	2	2	2	2	2	2
N	21996	21996	78208	78208	16276	16276	33071	33071
Border	3.837***	3.645***	1.929***	1.745***	3.800***	3.753***	2.211***	1.091***
	(31.58)	(32.34)	(13.17)	(11.43)	(15.68)	(15.88)	(8.19)	(3.46)
Gross output exp.	0.961***	0.951***	0.878***	0.888***	0.341***	0.341***	0.657***	0.658***
	(44.25)	(62.65)	(46.68)	(45.24)	(3.50)	(3.50)	(8.58)	(8.74)
GDP importer	0.680***	0.589***	0.673***	0.677***	0.976***	0.988***	0.978***	0.920***
	(13.28)	(11.03)	(20.22)	(22.36)	(18.09)	(16.90)	(19.29)	(18.04)
Distance	-0.486***	3.569**	-0.810***	-0.203	-1.254***	-2.455	-1.488***	1.730***
	(-6.12)	(2.23)	(-12.23)	(-0.84)	(-11.62)	(-1.67)	(-12.23)	(3.54)
Distance^2		-0.307**		-0.058***		0.083		-0.279***
		(-2.55)		(-2.70)		(0.80)		(-6.49)
Cont. Int. Part	0.149	0.275*	-0.234*	-0.350***	0.289	0.146	-0.208	-0.569**
	(1.14)	(1.77)	(-1.84)	(-2.60)	(1.37)	(0.57)	(-0.88)	(-2.37)
Cont. Nat. Part			0.156**	0.184**			-0.206	0.181
			(2.35)	(2.51)			(-1.36)	(1.31)
Pseudo R2	0.970	0.970	0.797	0.797	0.734	0.734	0.518	0.522

Note: All specifications include industry-specific exporter, industry-specific importer and year dummies. t-statistics in parenthesis with *** denoting significance at the 1-percent level, ** significance at the 5-percent level and * significance at 10-percent level.

Table 5. Robustness analysis III. Regions as origin

	Sample SPAIN				Sample REGIONS			
	(1)		(2)		(3)		(4)	
Origin								
Regions	15		15		15		15	
Destination								
Regions	-		-		15		15	
Rest of Spain	1		1		-		-	
Countries	17		17		17		17	
Sectors	13		13		13		13	
Years	2		2		2		2	
N	7020		7020		12480		12480	
Border	3.184	***	2.843	***	0.431		0.960	**
	(9.53)		(7.88)		(1.09)		(2.05)	
Gross output exp.	0.654	***	0.658	***	0.325	*	0.328	***
	(4.77)		(4.74)		(1.86)		(1.86)	
GDP importer	1.092	***	1.149	***	0.838	***	0.869	***
	(12.97)		(12.75)		(8.40)		(7.63)	
Distance	-1.414	***	-8.572	**	-2.077	***	-3.564	***
	(-6.68)		(-2.20)		(-14.88)		(-2.61)	
Distance^2			0.494	*			0.132	
			(1.88)				(1.18)	
Cont. Int. Part	-0.302		-0.836	***	-0.960	***	-0.776	**
	(-1.32)		(-3.34)		(-3.61)		(-3.27)	
Cont. Nat. Part					1.038	***	1.029	***
					(5.36)		(5.19)	
Pseudo R2	0.355		0.356		0.203		0.203	

Note: Tobit estimations, sample mean elasticities. All specifications include industry-specific exporter, industry-specific importer and year dummies. Regions-clustered robust t-statistics in parenthesis with *** denoting significance at the 1-percent level, ** significance at the 5-percent level and * significance at 10-percent level.