# The Gravity of Globalization

**Diego Agudelo and Larry Davidson**[1]

**August 2005**

*Diego Agudelo is Assistant professor at EAFIT University, Medellín, Colombia and is pursuing his doctoral degree in finance at the IU Kelley School of Business*

*Larry Davidson* is Professor of Business Economics and Public Policy*

*Kelley School of Business, Indiana University, Bloomington, IN 47405, USA*

*Abstract*

Can changes in the trade of the world's largest trading countries be considered more global? Or should they be labeled as more regional? We investigated these questions for the G7 countries for the time period from 1980 to 1997. We found that the usual dichotomy of global-regional is not rich enough to answer these questions because globalization can be measured in terms of both physical and cultural distance. Our new taxonomy allows for testing these separate impacts on world trade and suggests that trade changes are best described as *regional, though with some qualification*. With respect to physical distance, we find that trade is clearly becoming more regional. On the cultural dimension, however, we find conflicting results. These results are robust to a series of tests. We find the same pattern at industry level, except for Paper Products and Motor Vehicles. The regionalization pattern holds for both imports to and exports from the G7, but it is stronger for exports.

---

* Corresponding author. Kelley School of Business, Room 436, Bloomington, IN 47405-1701. Phone: (812) 855-2773. Fax: (812) 855-3354. davidso@indiana.edu

[1] The authors thank several colleagues without implication for their comments and encouragement: Michele Fratianni (the editor), Heejoon Kang, Alan Rugman, and Juergen Von Hagen. The authors thank for their financial support the Center for International Business Education and Research at the Indiana University Kelley School of Business.

## 1. Introduction

The central focus of the research is a test to measure changes in the degree of regionalization and globalization (heretofore, RZ and GZ respectively) for the G7 countries (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States) between 1980 and 1997. Rugman et al (2001, 2004, 2004a) challenged the notion that the post-war period is best described by GZ and provides instead evidence in favor of RZ. This raises several issues, including how to best define RZ and GZ as well as how to measure their changes over time. [2] Davidson (2002, 2004, 2004a) analyzed state-level U.S. export data to find evidence of both RZ and GZ. This paper suggests new definitions of these concepts and tests them with well-known gravity equations.

Our tests use Robert Feenstra's (1997) world trade flow data (bilateral import/export trade flow data for most countries, broken down by industry) and Andrew K. Rose's (2004) cultural, distance, and economic country data.[3] By estimating gravity equation parameters, we can measure the degree of change of RZ and GZ for the G7 countries. This approach uses the G7 group as a proxy for industrial countries and examines the trade of these seven countries with all their trading partners. This data set also allows us to probe further and examine these changes in GZ and RZ for the most important industries of the G7.

There is no single widely accepted definition of globalization in the international trade literature. To some people, globalization means a time period in which international trade increased at a faster pace[4]. At the other extreme is the idea that globalization is a new epoch – a time period that is qualitatively different from a previous period. While faster growth in international trade is one component of this larger view, it also contains the idea that trade goes above and beyond

what went before. It suggests that trade overcomes physical barriers and stretches over longer distances. Implicit in this view is that trade also transcends cultural barriers and that that traders go "farther" culturally by doing business in countries with new and different languages and religions. Under globalization, trade is enhanced not only by reductions in transportation and communications costs but it is also expanded by trends that make it more desirable and possible to engage persons from different cultural backgrounds. The world is "smaller" in terms of moving across language and religions as well as the ease of moving across physical distance. Thus, our tests of GZ/RZ go beyond tests of physical distance to examine cultural distance. If international business is increasingly conducted with countries that are physically farther (closer) – this would be one form of evidence in favor of GZ (RZ). If international business is increasingly done with countries that have dissimilar (similar) languages and religions, this would be further evidence of GZ (RZ). Thus, in our tests an unambiguous finding of GZ requires a decline in the effects of both physical and cultural distance on trade– a finding that companies have "scaled" the world's kilometers, languages, and religions.

These distinctions have implications. One might say that international trade always involves greater distance and cultural diversity. But this would be stretching the point. If Spanish companies decide to trade more with new business partners in its former colonies in South America, the knowledge requirements and other business challenges are likely to be significantly different from those involved with new international deals in China or Moldova. If one thinks we are in a New Age of globalization when in fact most of the new trade is regional (in terms of distance and/or culture), then business executives may be preparing themselves inappropriately.

We conclude that trade changes for the G7 countries between 1980 and 1997 are best described as RZ with respect to physical distance. That is, we find trade distances were declining, not increasing. We find mixed results with respect to cultural distance. Trade increased more with countries with different languages than with common-language partners (evidence of GZ). The result is just the opposite with respect to religion – trade was increasing more with countries of similar than dissimilar religions (evidence of RZ). These results remain statistically significant after performing a series of robustness tests. Most importantly, the economic effect of distance and language are substantial in all cases, while that of religion is economically small. These opposing cultural effects of religion and language suggest that Rauch's[5] network effects are at work, that is more trade with close countries that speak different languages. This higher language barrier might be made more scalable if trade agents sought out partners of the same religion. As a result we tried without positive results various immigration and foreign-born population numbers to investigate other determinants of common networks across countries. The language/religion results were unaffected by these additional tests.

The above results hold clearly for eight industries: Raw Materials, Non-electrical Machinery, Textile Products, Food and Related Products, Industrial Chemical, Ferrous industries, Household Audio-Video and Non-ferrous industries. Motor Vehicles, on the other hand, exhibited neither GZ nor RZ in the physical and in the language sense. Paper Products presented a clear pattern of GZ, but only in the imports to G7. The remaining eleven industries can be characterized as RZ in varying degrees and with increasing trade in different language countries.

## 2. Gravity Equations

We use gravity equations to estimate changes in GZ and RZ. Gravity equations have been used extensively in economic studies to estimate factors determining the size of flows (of capital, people, goods) between two geographic entities (cities, states, countries). More specifically, gravity equations have been employed recently to estimate the impact of currency unions and free trade agreements on international trade.[6] We know of no study that has estimated changes in GZ and RZ in the post-war period.

In its simplest form the gravity equation proposes (borrows heavily from hard science applications of the pull of gravity) that the flow of activity (trade) is proportional to the product of the "size" of the two entities and inversely proportional to the distance. In the case of international trade, we have for countries i and j:

(1) $\text{Trade}_{ij} = a\,(\text{GDP}_i \times \text{GDP}_j) / \text{distance}_{ij}$,

where GDP is a measure of economic size, distance is some measure of trade resistance, usually representing transportation and other costs related with the physical separation between the countries. A more general version of the gravity equation acknowledges the presence of information costs. Those costs are not only associated with physical distance but also with the cultural differences between the trade partners[7]. Accordingly, a log version of the gravity equation (1) can be written as follows:

(2) $\log \text{Trade}_{ijt} =$

$$a_0 + a_1 \log(\text{GDP}_{it} \times \text{GDP}_{jt}) + b_1 \log(\text{physical distance}_{ij}) + b_2\, \text{cult\_distance}_{ij} + \sum c_k Z^k_{ijt} + e_{ijt},$$

where physical distance$_{ij}$ is time invariant and is measured in miles or kilometers; cult_distance$_{ij}$ : measures the time-invariant cultural dissimilarity along the dimensions of language, religion, and migration between i and j[8]; $Z^k_{ijt}$ represents "k" control variables $Z^k$ usually incorporated in the bilateral trade gravity equation; and $e_{ijt}$ is a random error term with the usual properties.

The distance model predicts that $b_1$ and $b_2$ should be negative. While physical and cultural distance may be fixed over time, their impacts may not. A decline in the costs of either form of distance is tantamount to a decline in distance barriers and numerically smaller $b_1$ and $b_2$. If, however, the opposite holds, distance becomes a larger drag on trade and evidence in favor of RZ.

Making use of our full data set, we begin with the basic model, where t = 1 to 18 for the years from 1980 to 1997 and augment this equation to test for changes over time by adding interactive terms[9]:

$$(3)\ \log Trade_{ijt} = a_0 + a_1 \log (GDP_{it} \times GDP_{jt}) + b_1 \log(\text{physical distance}_{ij}) + b_{1a}\, t \times \log(\text{physical distance}_{ij}) + b_2\, \text{cult\_distance}_{ij} + b_{2a}\, t \times \text{cult\_distance}_{ij} + \sum c_k Z^k_{ijt} + e_{ijt}$$

The null Hypothesis is no change in the role of physical or cultural distance between 1980 and 1997:

$$(4)\ H_0 : b_{1a} = b_{2a} = 0$$

The alternative hypotheses are related to our definition of RZ and GZ as follows:

| Culture / Distance | Similar | Dissimilar |
|---|---|---|
| **Close** | **RZ**<br>$b_1 < 0$, $b_{1a} < 0$<br>$b_2 < 0$, $b_{2a} < 0$ | **Mixed**<br>$b_1 < 0$, $b_{1a} < 0$<br>$b_2 < 0$, $b_{2a} > 0$ |
| **Far** | **Mixed**<br>$b_1 < 0$, $b_{1a} > 0$<br>$b_2 < 0$, $b_{2a} < 0$ | **GZ**<br>$b_1 < 0$, $b_{1a} > 0$<br>$b_2 < 0$, $b_{2a} > 0$ |

In the empirical tests, cultural variables are defined in terms of cultural proximity (common language, common religion) rather than in terms of cultural distance. This only means that the expected signs of the cultural coefficients, $b_2$ and $b_{2a}$, are going to be the opposite to those indicated in the table above, without any loss of generality.

## 3. Regression Results

We begin by presenting results of the basic gravity equation applied to total trade in goods of the G7 counties and trade with 146 partners from 1980 to 1997. The list of countries included in this study is presented in Appendix B. We then add time interactive dummies to investigate changes in the impacts of key variables over time. Robustness tests are evaluated and we end the all industries part with a discussion of the economic significance of our results. A final set of results analyzes industry effects.

### 3.1 All industries, full time period

The left-hand-side variable is the log of real bilateral trade in US$ between each one of the G-7 countries and 146 trade partners, from 1980 trough 1997 (from the database of bilateral trade in US$ provided by Feenstra (1997)). Out of a total of 18.018 observations, 306 were dropped

because they had zero bilateral trade, leaving a total of 17,712 observations. The right-hand side variables, after Rose (2003), (see Appendix A for more information), were[10]:

| | |
|---|---|
| *Log_prod_gdp* | log of the product of the two real GDPs in 1995 US$ |
| *Logdistance* | log of distance, in miles |
| *Log_areas* | log of product of areas in squared miles |

Dummy variables accounting for :

| | |
|---|---|
| *Comlang* | Common language |
| *Colony* | Ever in a colony relationship |
| *Comcur* | In a strict currency union/ 1:1 peg |
| *RTA* | In the same free trade agreement |
| *Landlocked* | Number of landlocked countries in the pair (0,1,2) |

We began with a pooled regression with yearly dummy variables, to account for fixed effects of time. Fixed effects are pervasively used in panel data models to account for omitted year effects, e.g worldwide economic growth or decreasing cost of shipment. Therefore, in all the different specifications we include unreported yearly dummy effects. This first equation (Column A of Table 3.1) does not include time interactive variables, and include robust standard errors (clustering country pairs) following Rose (2003).

*[ Insert Table 3.1 Here]*.

All the estimated parameters are significant and with the right sign: the effects of *Log_prod_gdp*, *Colony*, *Comlang*, *Comcur*, and *RTA* are positive; the effects of *Landlocked*, *Log_areas* and *Logdistance* are negative; the $R^2$ is quite high (82%), and all the yearly dummy variables are quite significant (not reported).

To test for changes in GZ/RZ over time, we turn to an evaluation of the changing effects of distance and culture on trade from 1980 to 1997. Recalling that cultural variables are measured in terms of proximity, the expected signs on the coefficients of physical distance are opposite to those of cultural distance.

The key variable for physical distance is *Logdistance* and for cultural proximity are *Comlang* and *Colony*. Estimating cultural proximity is arguably more complex than a common language and/or colonial relationship. Thus, we add a religious proximity variable.[11]

Initially, we measured religious proximity by the percentage of people in each country affiliated with each of the major religion denominations --Catholic, Protestant, Orthodox, Jewish, Muslim, Buddhist, and Hindu--[12] Then, we calculated the first of our religious similarity variable (*Religprox*) as follows:

$$\begin{aligned}\text{Religprox}_{ij} &= \%\text{Cath}_i \times \%\text{Cath}_j + \%\text{Prot}_i \times \%\text{Prot}_j + \%\text{Ortod}_i \times \%\text{Ortod}_j \\ &+ \%\text{Jew}_i \times \%\text{Jew}_j + \%\text{Muslim}_i \times \%\text{Muslim}_j + \%\text{Bud}_i \times \%\text{Bud}_j + \%\text{Hind}_i \times \%\text{Hind}_j\end{aligned} \tag{5}$$

*Religprox* is higher the larger is the proportion of people from country i and country j that share the same religion. This variable can also be interpreted as the probability that a person, after a random draw from each country, may share the same religion.

After including this variable in the model, its regressor appeared with the expected positive sign, and was statistically significant, as shown in column B of Table 3.1. Consequently, we infer that

the religious proximity variable accounted for cultural dimensions not directly measured either by *Comlang* or by the *Colony* variables[13]

### 3.2 Time interactive effects

To investigate intertemporal changes of the effects of the cultural and distance variables we added time interactive variables in the model. The new empirical variables are then formed as the product between a trend variable t (= 0 in 1980, = 1 in 1981 and so) and the variables ldist, *Log_prod_gdp*, *Colony*, *Comlang* and *Religprox*.

We included the interactive effect of the log of the product of real GDPs, since this is the single most dominant variable in the model, explaining by itself 71% of the variance of ltrade[14]. We did not add time interactive dummy variables for fta, *Landlocked*, *Comcur*, and *Log_areas*..

Column C of Table 3.1, henceforth referred to as the "Base model", shows reinforcing effects of distance and religious proximity over time, both strongly statistically significant. There are also marginally significant effects (at 10%) of the log of GDP (increasing) and the common language (decreasing). There are no important multicollinearity problems, as indicated by a maximum and mean variance inflation ratio of 4.97 and 2.73, respectively.[15]

These findings are quite robust to different specifications. Following Rose (2003), we tried, alternatively, models with country-pair fixed and random effects to provide for potential omitted country-pair effects, a Prais-Winstein model with random effects to account for first-order autocorrelation of the residuals in the model, and a Tobit regression with random effects (Tobit

R.E.), as used by Chen (2003), which admits observations with zero trade. The results are qualitatively the same, and are available from the authors upon request.

### 3. 3 Robustness test, Country Exclusions

Next, we tested the model's robustness to country changes by excluding one of the G-7 countries at a time. Table 3.2 shows the results for excluding each country.

*[ Insert Table 3.2 Here]*.

In general, these findings for specific countries are consistent or at least not contrary to those obtained with the pooled data model, and provide additional information. The increasing negative effect of distance on trade is especially robust, and of similar magnitude after dropping any of the G-7 countries. The decreasing positive effect of common language seems concentrated mainly in the data of France, Italy, UK and USA[16]. The decreasing effect of the variable *Colony* is not robust to excluding any country, with the exception of USA. The increasing positive effect of common religion seems concentrated in Japan and in Italy. [17]

### 3.4 Further robustness tests

We submitted the model to a further series of robustness tests[18]. To verify that the results were not driven by a subset of very small or very poor countries, we considered excluding trade partners classified in the lowest three deciles of GDP in 1997 and, independently, excluding the countries in the three lowest deciles of trade for each G-7's in 1997. The excluded-small country versions are labeled B and D, for GDP and trade, respectively, in Table 3.3.

We also ran an "error-in-variable" model to account for the estimated imprecision in measuring real GDP. Since the log of the product of real GPD, *Log_prod_gdp*, is the dominant variable of the model, any significant imprecision in its measurement casts doubts on the robustness of our results. To estimate that imprecision we use the $R^2$ between the *Log_prod_gdp* used by Rose (2003) (calculated from GDP data from Penn World tables, WDI and IFS statistics) and the one used by us (mostly from WDI) obtaining a pooled correlation of about 91% for the common sample. Then, to account for a potential error in the estimation of this variable, we run the "error-in variable model" (model F in Table 3.3) with a reliability of the *Log_prod_gdp* value of 0.91.

*[ Insert Table 3.3 Here]*.

The increasing negative effect of the log of distance is robust in all the specifications, but not in those that focus on the smaller countries by GDP or trade. So these effects appear not to be driven by a small country effect. The increasing effect of the log of GDP over time doesn't fare well: It seems positive and strong when we focus only on the largest countries, but it changes signs when we account for the possible error in measuring the GDP. The decreasing effect of the common language over time is statistically significant in the relevant specifications. The increasing effect of the religious proximity variable is robust across the different specifications but in the last one, and it is clearly not driven by a small-country effect, on the contrary, it is stronger when we exclude the smaller countries either by GDP or by trade.

Finally, one might think that the mentioned time-varying effects of distance, common language and similar religion might be reflecting the increase in trade of the G7's with the countries that

abandoned communism in the early 90's, after the fall of the Berlin Wall[19]. In particular, for the four European G7 countries, the boost of trade with Eastern Europe in the 1990s seems clearly a case of increasing trading at shorter distance and with non-common language nations. To control for that, we ran the basic gravity equation dropping the observations of the ex-communist countries[20], as reported in the column B of Table 3.4. Besides, we ran model C with a dummy variable (*ex_com*) for the ex-communist trade partners, and model D with a trend variable starting in 1991 for the ex-communist trade partners (*ex_comm_trend*). It is expected that the estimator of *ex_com* be negative, since it should capture the incremental difficulty in trading with those nations, while the estimator of *ex_comm_trend* be positive, reflecting the gradual rising of trade with them upon the fall of the Berlin Wall.

*[ Insert Table 3.4 Here]*.

The resulting estimators of the variables *ex_com* and *ex_comm_trend* are highly significant and with the expected sign: the estimator of *ex_com* reflects the existence of barriers for the trade with the ex-communist countries, not accounted by the other variables of the gravity equation. The positive estimator of *ex_comm_trend* can be interpreted as a slow but continuous dismantling of those barriers. Adding these two variables significantly increases the R2 of the model We leave a more complete analysis of this Fall of Berlin Wall for future study.[21]

Regarding to the increasing effect of religious proximity, the results of Table 3.4 are unambiguous: None of this seems attributable to an increase in trade with ex-communist countries. In the case of the increasingly negative effect of distance, models B and D suggest that only a small part of it might be attributed to an increased trade with the ex-communist countries.

On the other hand, in models B and D the estimator of the interactive effect of common language and time, while still negative, loses statistical significance, although is still economically significant as presented in the next section. This seems to imply that a good part of the decreasing effect of common language in the G7 trade might be attributed to an increased trade with the ex-communist countries with respect to the rest of the world.

### 3.5 Economic significance

Taking the results of the Base Model (Table 3.1, column C), we can estimate the economic significance of the variables of interest. Overall, these estimations show that the increasing effect of distance, and the decreasing effect of common language over time are substantial, while the positive effect of religious proximity is modest.

First, the effect of *Comlang* in 1980 was exp(0.46)-1 or 58 %, That is, on average the group G-7 traded 58 percent more with similar-language countries than with dissimilar-language counties, ceteris paribus. This effect decreased at a yearly rate of exp(-0.0099)-1= -0.98%, so by 1997 the G-7 group traded only 34 percent more with similar-language countries. On the other hand, taking the results of column B of Table 3.4, after excluding the ex-communist countries, the decreasing effect of common language is still important going from 43% in 1980 to 29% in 1997.

The quantitative effect of distance can be shown as follows: in 1980 a 100 percent, ceteris paribus, increase in distance meant a exp(-0.676 × log2)-1 = -37% of change in trade; while the same effect was exp((-0.676+17 × -0.0118 ) × log2)-1= -45.5 % in 1997. [22]

Finally, going from 2% to 30% of religious similarity (interquartile range for the entire dataset) increased trade by (exp(0.00072× 28))-1 = 2.5% in the 1980. The same effect was of (exp(0.00072 +17×0.0017)× 28))-1 = 10.9% in 1997.

### 3.6 Industry effects

The database of Statistics Canada also provides the imports and exports between country pairs discriminated for the 34 industries defined by the BEA[23]. This allows us to explore to what extent the results from the gravity equation differ across industries. To do this we begin with an industry-level gravity equation based on the one proposed by Chen (2003). This model is similar to the country-level gravity equation (2) but includes industry specific variables, as follows:

$$(6)\quad \log(\text{Import}_{ijt,\,p}) = a_0 + a_1 \log(D_{it,\,p}) + a_2 \log(Y_{jt,\,p}) + b_1 \log(\text{distance}_{ij}) + b_2\, \text{cult\_proxim}_{ij}$$

$$\sum c_k Z^{k}_{ij} + \sum c_{k'} Z^{k'}_{ijt} + \sum \text{dummy year}_{\mathbf{t}},$$

where $\text{Import}_{ijt,\,p}$ are the imports in real dollars to country "i" from country "j" of goods classified in the industry “p”, in the year "t". At least one of the two “i" or “j” is a G7 country. We refer to this as “imports” only as a matter of convenience, since this variable represents both types of unilateral trade of the G7: imports and exports. The main explanatory variables are the log of $D_{it,\,p}$ the demand of products of industry “p” in country “i” in year “t”, the log of $Y_{jt,\,p}$ the production of industry “p” in country “j” in year “t”, the log of the distance, and the cultural proximity variable (*Comlang* or *Religprox*). Additionally, from the country-level model we include the time-invariant variables $Z^{k}_{ij}$ (*Log_areas*, *Landlocked* and *Colony*), time-dependent control variables $Z^{k'}_{ijt}$ (*Comcur* and fta) and year-dummy variables.

The industry gravity equation holds in any given time, so if we subtract the equation (6) for the period t1 from the one for the period t2 , with t2>t1, we obtain:

(7) $\log (Import_{ijt2,\,p} / Import_{ijt1,\,p}) = a_{1a} \log (D_{it2,\,p}/D_{it1,\,p}) + a_{2a} \log (Y_{jt2,\,p} / Y_{jt1,\,p}) + \sum c_{k'} \times (Z^{k'}_{ijt2} - Z^{k'}_{ijt1}) + a_{0a}$

This way we are explaining the growth of the imports as a function of the growth of the industry-specific demand for the importer, the growth of the industry-specific production of the exporter and changes in the time-dependent control variables. Note that the effects of the not-time-dependent variables are dropped, and that the effects of the year dummies are subsumed in the constant "$a_{0a}$" of the model. If we include the variables *Logdistance* and *Comlang* in (7) they should come up insignificant if and only if the effects of this two variables in the industry trade are fixed over time. As a consequence, any explanatory power that a time-invariant variable such as ldist, and *Comlang* may have in this model is just reflecting their time-varying effect. Then, model (7) is modified accordingly, to also include industry dummy variables to account for omitted industry-specific effects:

(8) $\log (Import_{ijt2,\,p} / Import_{ijt1,\,p}) =$

$$a_{1a} \log (D_{it2,\,p}/D_{it1,\,p}) + a_{2a} \log (Y_{jt2,\,p} / Y_{jt1,\,p}) + b_{1a} \log(\text{physical distance}_{ij}) + b_{2a}\, Comlang_{ij} + \sum c_{k'} \times (Z^{k'}_{ijt2} - Z^{k'}_{ijt1}) + \sum d_p \times (\text{industry dummy}_p) + a_{0a}$$

The null hypothesis is that the distance and cultural proximity variables provide no explanation to trade growth in the different industries ($b_{1a}$ =0, $b_{2a}$ =0). If those estimators turn out significantly positive (negative), that would imply increasing (decreasing) trade growth at longer distances or with language similarity.

The estimation of (8) requires several steps. First, since data on industry specific demand and production is not available for each of the 34 industries and the 147 countries included in this work, we proceed as follows[24]:

$$(9a) \quad D_{it,p} = GDP_i \times \sum_j Import_{ijt,p} \Big/ \sum_p \sum_j Import_{ijt,p} \qquad (9b) \quad Y_{jit,p} = GDP_j \times \sum_i Import_{ijt,p} \Big/ \sum_p \sum_i Import_{ijt,p}$$

Second, instead of taking any arbitrary initial and final years, we measure the growth of unilateral trade using five-year averages closest to the sample end-points: the average real trade from the period 80-84 and the average real trade from the period 93-97; we do the same for industry specific demand and production in (9a) and (9b). This way we are estimating the overall increase in trade over the entire period, using most of the data, while smoothing the effect of possible outliers[25]. Finally, to avoid obtaining results driven by a small-country or small-industry effect, for each industry we select the top observations representing 95% of the total bilateral trade. Out of a total of 47,067 industry country-pair observations, we end up with 14,903 covering 95 per cent of the trade for each particular industry.

The results of model (8) are presented in column A of Table 3.5, where we are pool together imports and exports of the G7 countries. The time-varying effects of the log of distance and the common language variable are both negative and highly significant. Moreover, the economic significance of those effects is quite similar to that reported for the all industries models: doubling the trading distance meant an average reduction of 9% on the growth of imports from 1980 to 1997, while trading with a common language partner meant an average reduction of 15% on the same variable. Table 3.5 also presents the results for imports-only and exports-only. The mixed-RZ pattern is present in both groups, but the magnitude of the RZ effect is twice as large for exports than for imports.

*[ Insert Table 3.5 Here]*.

Next, we investigate the different effects of distance and common language across industries. Model (8) can be easily extended to investigate this by replacing the distance variable with interactive variables between log of distance and the industry dummies, and similarly for the common language variable, as follows:

(10a) $\log (\text{Import}_{ijt2,\ p} / \text{Import}_{ijt1,\ p}) = a_{1a} \log (D_{it2,\ p}/D_{it1,\ p}) + a_{2a} \log (Y_{jt2,\ p} / Y_{jt1,\ p}) + b_{2a} \textit{Comlang}_{ij} + \sum c_{k'} \times (Z^{k'}_{ijt2} - Z^{k'}_{ijt1}) + \sum d_p \times (\text{industry dummy}_p) + \sum e_{1p} \times \textit{Logdistance}_{ij} \times (\text{industry dummy}_p) + a_{0a}$

(10b) $\log (\text{Import}_{ijt2,\ p} / \text{Import}_{ijt1,\ p}) = a_{1a} \log (D_{it2,\ p}/D_{it1,\ p}) + a_{2a} \log (Y_{jt2,\ p} / Y_{jt1,\ p}) + b_{1a} \textit{Logdistance}_{ij} + \sum c_{k'} \times (Z^{k'}_{ijt2} - Z^{k'}_{ijt1}) + \sum d_p \times (\text{industry dummy}_p) + \sum e_{2p} \times \textit{Comlang}_{ij} \times (\text{industry dummy}_p) + a_{0a}$

We estimate (10) alternatively by pooling imports and exports of the G7 countries and separating imports from exports.. The results of the interactive effect coefficients $e_{1p}$ and $e_{2p}$ are uninteresting by themselves; instead, we focus on their economic effects; see Table 3.6. For convenience, we only report the results of the top 20 largest industries for bilateral trade of the G7 and Raw Materials covering the period 1980-1984 and 92.5 per cent of total trade.

*[ Insert Table 3.6 Here]*.

The results of Table 3.6 can be summarized as follows:

- The RZ result, understood as a more negative effect of distance on trade growth, was largely driven by the following trade flows: The imports of Raw Materials; the exports of Non-electric machinery, Industrial Chemical and Ferrous industries; and the bilateral

trade of Textile products, Food and related products, Household Audio-Video and Non-ferrous industries.

- On the other hand the negative effect of common language on the growth of trade was due mainly to the following trade flows: the exports of Raw Materials and Textile products, the imports of Motor Vehicles and both the imports and exports of Non-electric Machinery, Industrial Chemicals and Other Transportation equipment.
- Most of the industries fall in the Mixed-RZ trend found in the all industry models. Only a few of the 21 industries showed trade patterns clearly opposed to the overall trend: First, the imports of Paper products showed a clear pattern of GZ: doubling the trading distance represents a 22% of increment of imports. Second, the imports of Raw materials, Textile products, Food and related products, Paper products and Electronic components showed sizable positive effects of common language, although none of them is statistically significant. Overall, the Mixed-RZ trend of the bilateral trade is strongly present in the Exports of the G7s but it is also present in the imports to the G7s, albeit to a lesser degree.
- A few of the top industries showed neither a trend for GZ nor for RZ in both the physical and the cultural dimensions. That's clearly the case for Motor Vehicles, the second largest industry , which doesn't present significant changes of growth due to distances or language, neither for imports nor for exports. That's also partially true for the industries of Instruments, Computers and Electronics, though the exports of those industries show trend for RZ.

## 4. Conclusions

Overall, the results indicate clearly that G-7 countries tended to trade over time with closer countries, countries with dissimilar languages and with similar religions.

The increasingly negative effect of distance is robust under and economically significant in all the different specifications considered. Although the common language decreasing effect is not statistically significant in some robustness tests, most importantly when the increasing trade with ex-communist countries is controlled, the estimators are always negative and economically significant. On the other hand, whereas the religion effect is robust to most specifications, it is economically quite small. All things considered, we believe that "regionalization with cultural qualifications" is the best way to describe the pattern of international trade for G7s in the sample period.

Focusing on the most important industries for trade in these countries suggests that most industries contributed to the general all-industries results. These results are present in both imports and exports to the G7 but are stronger for exports .

## References

Chen, Natalie (2003). Intra-national versus international trade in the European union: Why do national borders matter?, *Journal of International Economics*. Vol. 63, Pages 93-118

Davidson, Lawrence S. (2004). Regional Integration of U.S. Border States with Canada: Evidence from U.S. State Exports, in *North American Economic and Financial Integration*: Research in Global Strategic Management, Vol. 10, Oxford: Elsevier, 69-84.

Davidson, Lawrence S. (2004a). Was Europe a Balancing Force for the Regional Distribution of Exports from the United States? Evidence from U.S. State Exports, 1996 to 2001, in *Transatlantic Perspectives on Liberalization and Democratic Governance*, edited by David Audretsch and Charles Bonser, under review at LIT Verlag (Munster – Hamburg – Berlin – London), 58-77

Davidson, Lawrence S. (2002). Globalization and Changes in Industrial Concentration: State and Regional Exports from America's Heartland, chapter in *Globalization and Regionalization*, Challenges for Public Policy, edited by David Audretsch and Charles Bonser, Kluwer Academic Publisher, 49-70.

Egger, Peter (2002), "An Econometric View on the Estimation of Gravity Models and the Calculation of Trade Potentials," *World Economy*, Vol. 25, issue 2, pp 297-312.

Feenstra, Robert C., Robert E. Lipsey, and Harry P. Bowen (1997), "World Trade Flows, 1970 to 1992, with Production and Tariff Data," NBER Working Paper No. W 5910.

Feenstra, R. C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo.(2005), "World Trade Flows, 1962 to 2000," NBER Working Paper No. W 11040

Fratianni, Michele (2005), "Borders and Integration," this volume, chapter 2.

Frankel, Jeffrey and Andrew K. Rose (2002), "An Estimate of the Effect of Common Currencies on Trade and Income," *The Quarterly Journal of Economics*, May, 437-466.

Glick, Reuven and Andrew K. Rose (2002), "Does a currency union affect trade? The time-series evidence.," *European Economic Review*, June 2002, Vol. 46, issue 6, 1125-150

Hutchinson, William (2002), "Does Ease of Communication Increase Trade?" *Scottish Journal of Political Economy,* Vol. 49, 544-556.

Hutchinson, William (2002a), "The Gravity Equation in International Economics: Theory and Evidence," *Scottish Journal of Political Economy,* Vol. 49, 489-490.

Matyas, Laszlo (1998), "The gravity model: Some econometric considerations," *World Economy*, vol. 21, issue 3, 397-402

Mélitz, Jacques (2002). "Language and Foreign Trade," Discussion Paper Series. Centre for Economic Policy Research, No. 3590.

Otsubo, Shigeru T. and Tetsuo Umemura, Tetsuo (2003), "Forces Underlying Trade Integration in the APEC Region: A Gravity Model Analysis of Trade, FDI, and Complementarity," *Journal of Economic Integration*, vol. 18, no. 1, 126-49.

Rauch, James E. and Vitor Trindade, "Ethnic Chinese Networks in International Trade," (2002) *Review of Economics & Statistics*, Vol. 84, 116-130.

Rauch, James E. (2001) ,"Business and Social Networks in International Trade," *Journal of Economic Literature,* Vol. 39, 1177-1203.

Rose, Andrew K. (2000), "One money, one market: the effect of common currencies on trade," *Economic Policy*, vol. 15 issue 30, 9-46.

Rose, Andrew K. (2002), "Currency Unions and International Integration," *Journal of Money, Credit and Banking*, vol. 34, no. 4, 1067-89.

Rose, Andrew K. (2005), "Which International Institutions Promote International Trade?" *Review of International Economics*, Volume 13, 682-698.

Rose (2003), website for recent data, http://faculty.haas.berkeley.edu/arose/RecRes.htm, Nov. 2003.

Rugman, Alan (2001), *The End of Globalization*. London: Random House and New York: Amacom-McGraw Hill.

Rugman, Alan M. and Alain Verbeke (2004a). Regional Multinationals and Triad Strategy in Alan M. Rugman (Ed.) *Research in Global Strategic Management*: Volume 8, forthcoming.

Rugman, Alan M. and Alain Verbeke (2004). Regional and Global Strategies of Multinational Enterprises, *Journal of International Business Studies,* Vol. 35, 3 - 18.

Stulz, Rene M., and Rohan Williamson (2003), "Culture, Openness, and Finance," *Journal of Financial Economics*, Vol. 70, 313-349.

**Table 3.1 Gravity equations of the bilateral trade of G7 countries. 1980-1997**

| Variable | A | | B | | C | |
|---|---|---|---|---|---|---|
| *Landlocked* | -0.73833 | *** | -0.73829 | *** | -0.73866 | *** |
| | (0.08206) | | 0.08165 | | 0.08159 | |
| *Comlang* | 0.38518 | *** | 0.37456 | *** | 0.46008 | *** |
| | (0.09014) | | (0.08987) | | (0.10044) | |
| *Colony* | 1.31253 | *** | 1.3249 | *** | 1.41577 | *** |
| | (0.11549) | | (0.11565) | | (0.13519) | |
| *Comcur* | 0.72189 | ** | 0.6974 | ** | 0.70421 | ** |
| | (0.27436) | | (0.27625) | | (0.27664) | |
| *Log_areas* | -0.04805 | *** | -0.04804 | *** | -0.0485 | *** |
| | (0.01306) | | (0.01306) | | (0.01306) | |
| *Logdistance* | -0.7759 | *** | -0.76884 | *** | -0.67637 | *** |
| | (0.04897) | | (0.04846) | | (0.05069) | |
| *Log_prod_gdp* | 0.9054 | *** | 0.90617 | *** | 0.89159 | *** |
| | (0.01642) | | (0.01642) | | (0.0186) | |
| *RTA* | 0.21573 | ** | 0.2008 | * | 0.16013 | *** |
| | (0.10359) | | (0.10242) | | (0.10417) | |
| *Religprox* | | | 0.00217 | * | 0.00072 | |
| | | | (0.00121) | | (0.00149) | |
| *Comlang×t* | | | | | -0.0099 | * |
| | | | | | (0.00559) | |
| *Colony×t* | | | | | -0.01089 | |
| | | | | | (0.00754) | |
| *Logdistance×t* | | | | | -0.01181 | *** |
| | | | | | (0.00237) | |
| *Log_prod_gdp×t* | | | | | 0.00179 | * |
| | | | | | (0.00095) | |
| *Religprox×t* | | | | | 0.00017 | ** |
| | | | | | (0.00008) | |
| No. observations | 17712 | | 17712 | | 17712 | |
| $R^2$ | 0.82237 | | 0.82269 | | 0.82336 | |

This table presents the estimated coefficients of the log of bilateral trade by country-pair on several regressors. The sample consists of annual data spanning 1980–97 for the G7 countries and their trade partners. All the regressions include a constant and year dummy variables. *Logdistance* and *Log_prod_gdp* in the models B and C refer to the deviations of the mean of the log of distance and log of the product of the real GDP, respectively. Robust standard errors, calculated by clustering in country pairs, are shown below the corresponding coefficient estimates.
* Statistical significance at the 10% level. ** Statistical significance at the 5% level. *** Statistical significance at the 1% level

## Table 3.2 Country Exclusions

| Variable | exc_Canada | exc_France | Exc_Germany | exc_Italy | exc_Japan | exc_UK | exc_USA |
|---|---|---|---|---|---|---|---|
| *Landlocked* | -0.775*** | -0.747*** | -0.780*** | -0.713*** | -0.717*** | -0.773*** | -0.691*** |
| *Comlang* | 0.750*** | 0.405*** | 0.506*** | 0.479*** | 0.477*** | 0.472*** | 0.244** |
| *Colony* | 1.129*** | 1.421*** | 1.419*** | 1.406*** | 1.417*** | 1.397*** | 1.615*** |
| *Comcur* | 0.520** | 0.669** | 0.675** | 0.728** | 0.722** | 0.684** | 0 |
| *Log_areas* | -0.017 | -0.052*** | -0.053*** | -0.056*** | -0.057*** | -0.048*** | -0.046*** |
| *Logdistance* | -0.654*** | -0.687*** | -0.703*** | -0.634*** | -0.656*** | -0.698*** | -0.702*** |
| *Log_prod_gdp* | 0.836*** | 0.898*** | 0.900*** | 0.905*** | 0.907*** | 0.906*** | 0.880*** |
| *RTA* | 0.230** | 0.178 | 0.134 | 0.163 | 0.161 | 0.119 | 0.168* |
| *Religprox* | 0 | 0.002 | 0.001 | 0.001 | 0.001 | 0 | 0 |
| *Comlang×t* | -0.012*** | -0.012*** | -0.012*** | -0.012*** | -0.009*** | -0.014*** | -0.011*** |
| *Colony×t* | 0.002** | 0.002* | 0.001 | 0.002* | 0.003** | 0.002 | 0.002 |
| *Logdistance×t* | -0.004 | -0.014 | -0.012 | -0.01 | -0.011 | -0.005 | -0.016* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Log_prod_gdp×t* | -0.017** | -0.007 | -0.011* | -0.007 | -0.012** | -0.01 | -0.006 |
| *Religprox×t* | 0.00020193** | 0.00019201* | 0.00019506** | 0.00013442 | 0.00007013 | 0.00021851** | 0.00016414* |
| No. observations | 15160 | 15122 | 15120 | 15120 | 15120 | 15120 | 15132 |
| $R^2$ | 0.82126 | 0.82036 | 0.81039 | 0.81756 | 0.83124 | 0.8184 | 0.83288 |

This table presents the estimated coefficients of the log of bilateral trade by country-pair on several regressors The sample consists of annual data spanning 1980–97 for each of the G7 countries and their trade partners, excluding one of the G7s in each model. All the regressions include a constant and year dummy variables. Robust standard errors were calculated clustering by country pairs for all the models (not reported).
* Statistical significance at the 10% level. ** Statistical significance at the 5% level. *** Statistical significance at the 1% level

## Table 3.3. Further Robustness Tests

| Variable | A Base Model | B Excluding small_GDP | C Only small_GDP | D Excluding small_trade | E Only small_trade | F error-in-var model |
|---|---|---|---|---|---|---|
| *Landlocked* | -0.739*** | -0.520*** | -0.958*** | -0.569*** | -0.552*** | -0.478*** |
| *Comlang* | 0.460*** | 0.498*** | 0.350* | 0.501*** | 0.198 | 0.906*** |
| *Colony* | 1.416*** | 1.130*** | 1.810*** | 1.122*** | 1.701*** | 1.724*** |
| *Comcur* | 0.704** | 0.718*** | 0.455 | 0.49 | 1.762*** | 0.647*** |
| *Log_areas* | -0.049*** | -0.052*** | -0.022 | -0.039*** | -0.051** | -0.121*** |
| *Logdistance* | -0.676*** | -0.629*** | -0.999*** | -0.593*** | -1.090*** | -0.322*** |
| *Log_prod_gdp* | 0.892*** | 0.842*** | 1.061*** | 0.774*** | 0.989*** | 1.472*** |
| *RTA* | 0.16 | 0.152 | 0.548** | 0.221** | 0.954*** | 0.028 |
| *Religprox* | 0.001 | -0.003** | 0.011*** | -0.001 | 0.006* | 0.002*** |
| *Comlang×t* | -0.012*** | -0.010*** | -0.011 | -0.008*** | 0.006 | -0.039*** |
| *Colony×t* | 0.002* | 0.006*** | -0.016*** | 0.002** | -0.017*** | -0.044*** |
| *Logdistance×t* | -0.011 | -0.002 | -0.021* | -0.009 | -0.023 | -0.046*** |
| *Log_prod_gdp×t* | -0.010* | -0.014** | -0.004 | -0.016*** | -0.008 | -0.041*** |
| *Religprox×t* | 0.00017441** | 0.00032537*** | -0.00036873 | 0.00020428** | -0.00013848 | 0.00004645 |
| No. observations | 17712 | 12343 | 5369 | 12459 | 5253 | 17712 |
| $R^2$ | 0.82336 | 0.79505 | 0.57105 | 0.78375 | 0.51499 | 0.92509 |

This table presents the estimated coefficients of the log of bilateral trade by country-pair on several regressors. The sample consists of annual data spanning 1980–97 for each of the G7 countries and their trade partners. All the regressions include a constant and year dummy variables. Robust standard errors were calculated clustering by country pairs for all the models but D (not reported) .
* Statistical significance at the 10% level. ** Statistical significance at the 5% level. *** Statistical significance at the 1% level

## Table 3.4 Controlling for the effect of the fall of the Berlin Wall.

| Variable | A (Base Model) | B | C | D |
|---|---|---|---|---|
| *Landlocked* | -0.73866*** | -0.73579*** | -0.67418*** | -0.67376*** |
| *Comlang* | 0.46008*** | 0.35574*** | 0.39572*** | 0.36917*** |
| *Colony* | 1.41577*** | 1.40954*** | 1.40543*** | 1.39241*** |

| | | | | |
|---|---|---|---|---|
| *Comcur* | 0.70421** | 0.63132** | 0.62726** | 0.61369** |
| *Log_areas* | -0.04850*** | -0.04343*** | -0.04136*** | -0.04201*** |
| *Logdistance* | -0.67637*** | -0.77451*** | -0.76751*** | -0.80580*** |
| *Log_prod_gdp* | 0.89159*** | 0.89489*** | 0.89855*** | 0.90191*** |
| *RTA* | 0.16013 | 0.01756 | 0.03868 | -0.00169 |
| *Religprox* | 0.00072 | -0.00068 | -0.00024 | -0.0007 |
| *Comlang×t* | -0.01181*** | -0.00835*** | -0.01295*** | -0.00933*** |
| *Colony×t* | 0.00179* | 0.00188* | 0.00163* | 0.00138 |
| *Logdistance×t* | -0.01089 | -0.0111 | -0.01143 | -0.01007 |
| *Log_prod_gdp×t* | -0.00990* | -0.00574 | -0.01002* | -0.00673 |
| *Religprox×t* | 0.00017** | 0.00023*** | 0.00018** | 0.00024*** |
| | | | | |
| *ex_com* | | | -0.96845*** | -1.18890*** |
| *ex_com_trend* | | | | 0.13902*** |
| | | | | |
| No. observations | 17712 | 16460 | 17712 | 17712 |
| $R^2$ | 0.82336 | 0.83235 | 0.83164 | 0.8326 |

This table presents the estimated coefficients of the log of bilateral trade by country pair on several regressors. The sample consists of annual data spanning 1980–97 for each of the G7 countries and their trade partners. All the regressions include year dummy variables. Robust standard errors were calculated clustering by country pairs for all the models (not reported).

* Statistical significance at the 10% level. ** Statistical significance at the 5% level. *** Statistical significance at the 1% level

## Table 3.5 Industry model .

| Variable | A | B | C |
|---|---|---|---|
| *log_growth* $D_{ip}$ | 1.031*** | 0.807*** | 1.050*** |
| *log growth* $Y_{jp}$ | 0.781*** | 0.741*** | 0.704*** |
| *Logdistance* | -0.129*** | -0.076*** | -0.158*** |
| *Comlang* | -0.167*** | -0.184*** | -0.175*** |
| Δ*_RTA* | 0.633*** | 0.531*** | 0.716*** |
| Δ*_Curcol* | -4.240*** | -5.270*** | -3.048*** |
| Δ*_Comcur* | -0.331*** | -0.32 | -0.373*** |
| | | | |
| No. observations | 14903 | 7752 | 7151 |
| $R^2$ | 0.371 | 0.32176 | 0.42628 |

This table presents the estimated coefficients of the log of unilateral trade by country-pair and industry on several regressors. The sample consists of the average industrial data for the periods 1980-1984 and 1993-1997 for each of the G7 countries and their trade partners. All the regressions include industry dummy variables not shown. Model A is for the pooled dataset of imports and exports, model B for imports to the G7, and model C for exports to the G7. Robust standard errors not reported.

* Statistical significance at the 10% level. ** Statistical significance at the 5% level. *** Statistical significance at the 1% level

**Table 3.6 Economic effects of distance and common language on industry growth of trade 1980-84 and 1993-97.**

| Industry | BEA | % Share Total Trade 80-84 | Distance effect | | | Common language effect | | |
|---|---|---|---|---|---|---|---|---|
| | | | Pooled regression | Imports to G7 | Exports to G7 | Pooled regression | Imports to G7 | Exports to G7 |
| Raw Materials | | 25.1% | **-14%** | **-11%** | -8% | 0% | 33% | -27% |
| Motor Vehicles | 28 | 11.0% | -2% | 4% | -4% | -2% | -15% | -2% |
| Other Non-electric machinery | 23 | 6.7% | **-11%** | -6% | **-8%** | **-22%** | **-29%** | **-17%** |
| Industrial Chemicals | 13 | 5.7% | -2% | 5% | **-6%** | **-20%** | -19% | **-21%** |
| Textile product | 5 | 4.6% | **-12%** | **-4%** | **-11%** | -1% | 16% | -20% |
| Food and related product | 4 | 4.0% | **-13%** | **-9%** | **-9%** | 6% | 12% | -3% |
| Other Transportation | 29 | 4.0% | -10% | -7% | -5% | **-29%** | **-41%** | -20% |
| Ferrous Industries | 17 | 3.6% | **-17%** | -7% | **-17%** | -18% | -25% | -13% |
| Household Audio-Video | 25 | 3.5% | **-17%** | **-9%** | **-12%** | -11% | -18% | -10% |
| Nonferrous Industries | 18 | 3.3% | **-22%** | **-20%** | **-10%** | -1% | 2% | -9% |
| Instruments | 33 | 3.2% | **-10%** | -3% | **-9%** | -8% | -13% | -9% |
| Computer equipment. | 22 | 2.6% | **-12%** | -4% | **-9%** | -11% | -14% | -7% |
| Other electric Machinery | 27 | 2.1% | **-17%** | **-9%** | **-14%** | **-21%** | -27% | -16% |
| Other Manufacturing | 34 | 2.0% | **-15%** | **-9%** | **-10%** | -11% | -9% | -19% |
| Construction | 21 | 2.0% | **-12%** | -6% | **-10%** | **-38%** | **-57%** | -16% |
| Paper product | 7 | 1.9% | 11% | **22%** | **-7%** | 6% | 20% | -8% |
| Fabricated Metal Products | 19 | 1.9% | -5% | 2% | **-8%** | -21% | **-35%** | -8% |
| Wood, Furniture | 30 | 1.8% | **-15%** | **-4%** | **-17%** | **-30%** | -22% | -36% |
| Other Chemicals | 14 | 1.3% | **-8%** | **-5%** | **-5%** | **-26%** | **-34%** | -17% |
| Leather product | 6 | 1.1% | **-11%** | **-12%** | -2% | **-25%** | -16% | **-37%** |
| Electronic Components | 26 | 1.1% | -9% | 1% | **-13%** | 22% | 35% | 8% |

Economic effect of distance and common language obtained from estimations of models (10a) and (10b). Economic distance effects refers to the effect of doubling trade distance on the growth of unilateral trade (import, exports or both), estimated as : $2 \hat{} e_{1p}-1$
Economics effects of common language refers to the incremental growth of unilateral trade from trading with common language partners, compared with trading with non-common language partners. Estimated as $\exp(e_{2p})-1$.
Economics effects derived from coefficients significantly different from zero at the 5% level are **indicated in bold**.

-

## Appendix A . Definition of the variables in the gravity equation for Chapters 5 and 6

| Mnemonic | Definition and Source |
|---|---|
| *Ltrade* | Log of real trade in US$ of 1995 for the period 1980-1997 (this chapter), and for 1980-2001 (chapter 6th)<br>From two databases edited by Canada Statistics-Center for International data: for the years 1980 -1984 we used the CD-Rom "World trade flows, 1980-1997", while for 1985-2001 we used the database "World trade flows, 1985-2001" . As explained in Feenstra et al (1997) and Feenstra et al (2005) those databases compile import and export flows for 164 countries in nominal dollars based on original data, reported to the UN. The UN data have been reviewed by Canada Statistics to account for omissions, discrepancies and inconsistencies. To obtain the bilateral trade value we added the import and export flows for each single trading pair in each year. The bilateral trade data in current dollars was translated to real US$ 95 dollars using the U.S. Consumer Inflation index from the WDI database[26] |
| *Log_prod_gdp* | Log of the product of real GDPs, in 1995 dollars of the two trade partners. primarily from the WDI database |
| *Log_prod_gdppc* | Log of the product of real GDPs per cap in 1995 dollars of the two trade partners, based on the GDP and population taken from the WDI database |
| *Logdistance* | Log of distance in miles. Distance was calculated between the mass centers of the trade partners using the long-circle formula. The mass center geographical locations have been taken from the CIA World factbook 2005 [27] |
| *Island* | Number of islands in the trade pair =2, 1 or 0. |
| *Log_areas* | Log of the product of the areas in square miles. The areas have been taken from the CIA World factbook 2005 |
| *Border* | Dummy variable, =1 if the two countries share a common border, =0 otherwise. Borders as given in the CIA World factbook 2005 |
| *Comlang* | Dummy variable, =1 if the two countries have the same main language, =0 otherwise. Source CIA World factbook 2005 |
| *Landlocked* | Dummy variable, =0, 1, 2 depending how many countries in the trade pair are Landlocked. |
| *Colony* | Dummy variable, =1 if the two countries were ever involved in a colonial relationship with one another, =0 otherwise. Source Microsoft Encarta 2004. www.wikipedia.org. March 2005. |
| *Curcol* | Dummy variable, =1 if the two countries were in a colonial relationship in the year of the observation, =0 otherwise. Source Microsoft Encarta 2004. www.wikipedia.org. March 2005. |
| *Comcol* | Dummy variable, =1 if the two countries have the same colonizer =0 otherwise. Source Microsoft Encarta 2004. www.wikipedia.org. March 2005. |
| *Comcur* | Dummy variable, =1 if the two countries have the same currency in a particular year observation, =0 otherwise. Source Table 1 of chapter 2, this volume. Others from www.wikipedia.org. March 2005. |
| *RTA* | Dummy variable, =1 if the two countries were in a regional trade agreement in the year of the observation, =0 otherwise. Based on list of RTAs on www.wto.org/english/tratop_e/region_e/region_e.htm. |
| *t* | Time trend variable, =0 for observations in 1980, =1 for 1981 and so on, up to 21 for 2001 |
| *ex_com* | =1 when the trade partner is a country that ended communist rule in early 90's: Albania, Bulgaria, Former Czechoslovak, Former USSR, Former Yugoslavia, Hungary, Mongolia, Poland, Romania, and Vietnam. =0 otherwise. |
| *ex_com_trend* | =1, 2, 3, ….from 1991 on , =0 otherwise, when the trade partner is a country that ended communist rule in early 90's : Albania, Bulgaria, Former Czechoslovak, Former USSR, Former Yugoslavia, Hungary, Mongolia, Poland, Romania, and Vietnam.. = 0 otherwise |
| *log_growth* $D_{ip}$ | Log of 1+ rate of growth of the demand in country "i" for goods in industry "p", from 1980-84 to 1993-1997, as given by (9a) |
| *log growth* $Y_{ip}$ | Log of 1+ rate of growth of the production of industry "p" in country "j" from 1980-84 to 1993-1997, as given by (9b) |
| *SA* | Dummy variables =1 if either of the two countries is in South America (Excluding Guyanas), |
| *Log_prod_gdp_sa Logdistance_sa, Comlang_sa, Border_sa, Comcol_sa, Colony_sa, RTA_sa* | Interactive variables of SA variable with *Log_prod_gdp, Logdistance, Comlang, Border, Comcol,Colony and RTA* respectively with South America. For example:<br>*Log_prod_gdp_sa = Log_prod_gdp* if either of the two countries is in South America (Excluding Guyanas),<br>=0 otherwise. |
| *Log_prod_gdpxt, Logdistancext, Comlangxt, Borderxt, Comcolxt, Colonyxt,* | Interactive variables of the time variable *t* with *Log_prod_gdp, Logdistance, Comlang, Border, Comcol, and Colony,* respectively.<br>To avoid multicollinearity problems, *Log_prod_gdpxt, Logdistancext* were calculated based on the demeaned values of *Log_prod_gdp* and *Logdistance* respectively |

## Appendix B. Countries included in the study ( Chapters 5 and 6 )

**North America**
Bermuda
Canada
Mexico*
USA

**Central America and Caribbean Islands**
Bahamas*
Barbados*
Belize*
Costa Rica*
Cuba*
Dominican Rp.*
El Salvador*
Guatemala*
Guyana*
Haiti*
Honduras*
Jamaica*
Neth. Antilles
Nicaragua*
Panama*
St Kitts and Nev.*
Suriname*
Trinidad and Tobago*

**South America**
Argentina*
Bolivia*
Brazil*
Chile*
Colombia*
Ecuador*
Paraguay*
Peru*
Uruguay*
Venezuela*

**Middle East and North Africa**
Algeria*
Bahrain
Cyprus
Egypt*
Iran*
Israel
Jordan*
Kuwait
Lebanon*
Libya
Morocco*
Oman*
Qatar
Saudi Arabia
Syria*
Tunisia*
Turkey*
United Arab Emirates

**Sub-Saharan Africa**
Angola*
Benin*
Burkina Faso*
Burundi*
Cameroon*
Central African Rep.*
Chad*
Comoros*
Congo*
Congo, Dem. Rep.*
Djibouti*
Eq. Guinea*
Ethiopia*
Gabon*
Gambia*
Ghana*
Guinea*
Guinea-Bissau*
Ivory Coast*
Kenya*
Liberia*
Madagascar*
Malawi*
Mali*
Mauritania*
Mauritius*
Mozambique*
Niger*
Nigeria*
Rwanda*
Senegal*
Seychelles*
Sierra Leone*
South Africa*
Sudan*
Tanzania*
Togo*
Uganda*
Zambia*
Zimbabwe*

**Australia and Pacific Islands**
Australia
Fiji*
Kiribati*
New Caledonia
New Zealand
Papua N. Guinea*
Solomon Islands*

**Eastern Asia**
China*
Hong Kong
Japan
Korea Republic
Mongolia*
Taiwan

**South East Asia**
Brunei
Indonesia*
Laos *
Malaysia*
Philippines*
Singapore
Thailand*
Vietnam*
South Asia
Bangladesh*
Bhutan*
India*
Maldives*
Myanmar*
Nepal*
Pakistan*
Sri Lanka*

**Western Europe**
Austria
Belgium-Luxemburg
Denmark
Finland
France
Germany
Greece
Iceland
Ireland
Italy
Malta
Netherlands
Norway
Portugal
Spain
Sweden
Switzerland
UK

**Eastern Europe**[28]
Albania*
Bulgaria*
Former USSR*
Hungary*
Poland*
Romania*

Only for Chapter 5
Former Yugoslavia*
Former Czechoslovak*

* Developing nations as identified in the Global development Indicators database of the World Bank.

[2] If changes in world trade are truly best understood under the umbrella of RZ as Rugman suggests, we seek to ask why by more closely examining trade by country and industry. We wish there were good data sets to facilitate this research that would allow investigation into the full post-war time period, but no such data exists. Since there were major changes in world financial arrangements in the 1970s, we believe starting in 1980 has merit on its own.

[3] From http://faculty.haas.berkeley.edu/arose/RecRes.htm, Nov. 2003. See Appendix A for more information about data sources. We chose Feenstra's data over a similar database by Rose because Feenstra's had industry disaggregation of the trade flows. We perform a comparison of the Feenstra flows to those published by the Organization for Economic Cooperation Development. Finding many differences between the two data sets but the overall impression for tests like ours is that Feenstra's data is quite compatible with the OECD's (details available upon request). Because the databases do have differences it is possible that one could obtain different test results using another database.

[4] Rugman (2001) reviews several broader definitions of globalization and finally settles on the following: "the activities of multinational enterprises engaged in foreign direct investment and the development of business networks to create value across national boundaries." Rugman goes on (page 12) to say that while globalization might exist in a few sectors (consumer electronics), "…it never really occurred; it is a myth. Instead the vast majority of MNE manufacturing and service activity is (and always has been) organized regionally, not globally."

[5] Rauch (2001, 2002) found that business and social networks are used by companies in international trade to overcome informal trade barriers (weak contract enforcement and inadequate information). Rauch cites examples of Indian and Chinese networks operating at great distance.

[6] Rose (2000) used a data set composed of 22,948 pairs, from 186 countries, for 1970 to 1990 in 5-year intervals. He ran both a pooled regression and separated regressions for each year, obtaining $R^2$ between 0.57 and 0.72, and all coefficients with the right sign and significant: specifically, with the exception of $\beta_3$ (distance) and $\delta$ (bilateral exchange rate volatility) all the coefficients are positive. His analysis is basically cross sectional. Glick and Rose employed (2002) a very similar model, but in a panel data setting, providing for fixed and random effects alternatives, that control for the variation in the effect of common currency through time. Additionally, they included three new control variables: "$Area_iArea_j$" as the product of the two land masses, "Landl" the number of the land locked countries in the pair (0, 1 or 2), and "Island" being the number of island nations in the pair. Egger (2002) pointed out several problems of the traditional OLS cross-section approach in the gravity equations, being the most significant, not properly accounting for the effect of time in the model, heteroskedasticity, and autocorrelation. He proposed a model that included four control variables which reflect a country's freedom with respect to international exchange. He also included the real bilateral exchange rate in the model. He didn't use as many control variables as Rose (2002) and his definitions of production, size, and per-capita effects also differed somewhat from Rose's. Egger tested his model with a data set of exports between OECD countries and 10 central and eastern European countries over the period 1986-1997. He obtained very high $R^2$ and ran several robustness tests. There are at least two papers that test the impact of language on trade (Hutchinson (2002), Melitz (2002) ). Egger (2002) included variables for the contractual and legal environment that seemed relevant in his tests. Rose (2005) used a standard gravity equation with panel data covering 50 years and 175 countries to examine the effects of various international organizations (World Trade Organization, International Monetary Fund, and the Organization for Economic Cooperation and Development) on trade.

[7] Doing business at increasingly physical but not cultural distance implies needs for transportation communications and logistical information. If people are better informed about close-by events, increasing the physical distance of trade requires new sources of information about a wider variety of subjects. If instead, businesses are taking place at increased cultural distance, there will be additional requirements in terms of language, business practices and other aspects of culture.

[8] While physical distance between two countries is fixed between times (unless the legal borders change) the cultural proximity probably does change over time. We consider such changes to be small enough to ignore for our purposes. We admit, however, that an interesting extension of this work would treat trade and cultural proximity as mutually determined variables.

[9] Our tests allow for several other interaction terms to allow for inter-temporal parameter shifts of selected other right-hand-side variables.

[10] Initially the model included the variables *Island, Log_prod_gdppc, Curcol* and *Border* (see definition in Appendix A), included in the model of Rose (2004). However the first three variables turned out statistically insignificant, which might be due to the fact that we have a different and smaller data set – focusing only on the bilateral trade of the G7 countries. Besides, the effect of border was found negative, whereas in Rose's model the estimator of a shared border is positive and significant. Again, it might be the result of our focus on the G7 that produces the unexpected result. Having a common border doesn't add to the explanation of bilateral trade beyond what is already accounted by the variable *Logdistance*. Indeed, when we drop from the dataset either the observations for Canada, or Germany and run the model without these countries (unreported), the perverse effect of border disappears as if it were concentrated in one or both of them.

[11] Religion was also considered by Stulz and Williamsom (2003), who perform cross sectional comparisons of financial systems across the world.

[12] Sources: CIA World Factbook 04, Windows Encarta, and www.adherents.com., Nov. 2003.

[13] We also considered three additional cultural variables, that were discarded in favor of *Religprox*: 1)religprox2 was calculated after grouping Catholic, Protestant, Orthodox and Jewish under the name of "Judeo-Christian" and using a similar expression as the one presented above. 2) commainrelig", is defined as 1 if the two countries shared the same majority religion and 0 otherwise. 3) comcultreg, being 1 if the two countries share the same cultural region and 0 otherwise, as given by the Geography textbooks. None of those variables fared better than *Religprox* in the model.

[14] In so doing we found it necessary to replace the variables *Logdistance* and *Log_prod_gdp* with their deviations from the respective sample mean to avoid multicollinearity problems, and recalculate *Logdistance*$\times t$ and *Log_prod_gdp*$\times t$ accordingly. doing so does not change the estimators of interest in this study.

[15] We include also a $t^2$ term as explanatory variable and interactive variables between $t^2$ and the distance and cultural proximity variables, to account for possible nonlinearities in the effect of time (not reported). The interactive effects with $t^2$ turned out to be no significant and the numerical estimators of the interactive effects with time were virtually unchanged. We thank Juergen Von Hagen for this suggestion.

[16] However, in all the cases the estimators of *Comlang*$\times t$ are negative, and the decreasing effect of common language is economically significant in each of the models of Table 3.2 using the reasoning of section 5.5 "Economic significance" (not reported).

[17] Since the large increase of bilateral trade between USA and Canada, and USA and Mexico was one the most important facts in the last 20 years we run models excluding alternatively the data for USA and Canada, USA and Mexico, and USA and both countries (unreported). The results remain qualitatively the same, and quantitatively almost unchanged. We thank Alan Rugman who suggested to check for this. Additionally, to see whether the results are mainly driven by inter G7 trade we also run the model excluding the bilateral trade between G-7 countries (unreported). The inter G7 trade accounts for roughly 30% of the total trade of the G7 countries. The results are qualitatively the same, and quantitatively almost unchanged.

18 In other non-reported robustness tests the time interactive effects were calculated over a six-year horizon: we created, for each of the variables of interest (*Logdistance, Log_prod_gdp*, *Comlang, Religprox,* and *Colony*) one interactive variable for the period 1980-1986 and another for the period 1991-1997 . This way we were effectively investigating the different effects of each variable in each of the three six-year periods. In a different robustness test we used a stricter definition for the common language variable than the one used by Rose(2003). The results of those models are qualitatively the same of the Base Model and are available from the authors upon request.

[19] We thank Juergen von Hagen who suggested to check for this.

[20] In our sample those countries are: Albania, Bulgaria, Former Czechoslovak, Former USSR, Former Yugoslavia, Hungary, Mongolia, Poland, Romania, and Vietnam

[21] Interestingly, the result of model D suggest that, after controlling for all the other factors, G7 countries traded 70% less with a ex-communist country than with a non-communist country before 1990, but that this difference dropped to 19 % in 1997.

[22] The 100 percent of increment on distance can be justified in this analysis since the ratio between the third and the first quintiles of distance of the entire dataset is 2.18.

[23] The Bureau of Economic Analysis (BEA) classifies industries in 34 groups. The remaining part of the trade that doesn't belong to any industry can be identified as "Raw materials" (e.g. Vegetables, grains, livestock, oil, mineral products).".

[24] We thank the editor for this suggestion

[25] On top of that , not including the observations of the years 1985-1992 has the virtue of reducing, or perhaps eliminating, the effect on (8) of the autocorrelation of the residuals expected in model (7)

[26] World Development Indicators, from the World Bank. http://devdata.worldbank.org/dataonline/. February 2005. When not available for a particular country, values were taken from IMF's World Economic Outlook

[27]Taken from CIA World Factbook site http://www.cia.gov/cia/publications/factbook, on June 2005.

[28] The denominations 'Former' USSR, Czechoslovak and Yugoslavia group the current nations that constituted those entities. We didn't have data on trade for Former Czechoslovak and Yugoslavia republics for 1999-20001, so those nations were dropped for the study of Chapter 6.