KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY
Department of Business Economics and Public Policy

Handbook of International Business, 2nd ed., Oxford University Press

The Gravity Equation in International Trade

Michele Fratianni*

Abstract
This chapter offers a selective survey of the gravity equation (GE) in international trade. This equation started in the Sixties as a purely empirical proposition to explain bilateral trade flows, without little or no theoretical underpinnings. At the end of the Seventies, the GE was "legitimized" by a series of theoretical articles that demonstrated that the basic GE form was consistent with various models of trade flows. Empirical applications of GE expanded to cover a variety of issues, such as the impact of regional trade agreements, national borders and currency unions on trade, as well as the use of the equation to sort out the relative merit of alternative trade theories. A new wave of studies is now concentrating on the general equilibrium properties of the GE and finer econometrics points. The renewed interest of the academic profession in the development of the GE is undoubtedly driven by the equation's empirical success.

Keywords: gravity equation, trade theories, borders, regional trade agreement, currency unions.

JEL Classification: E58, F15, F33, G15

Draft date: August, 2007

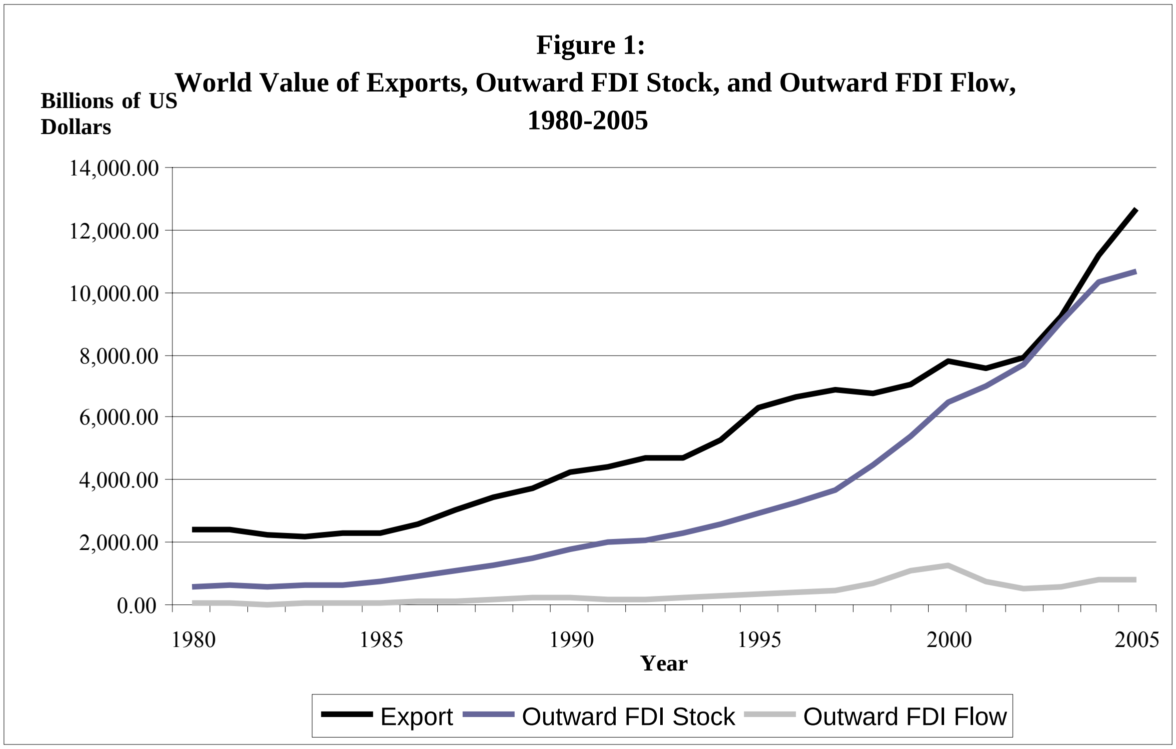
---

* Indiana University, Kelley School of Business, CIBER, Bloomington, Indiana 47405, USA and Università Politecnica delle Marche, Dipartimento di Economia, Piazza Martelli, 60121, Ancona, Italy. I thank Chang Hoon Oh for comments and very competent research assistance.

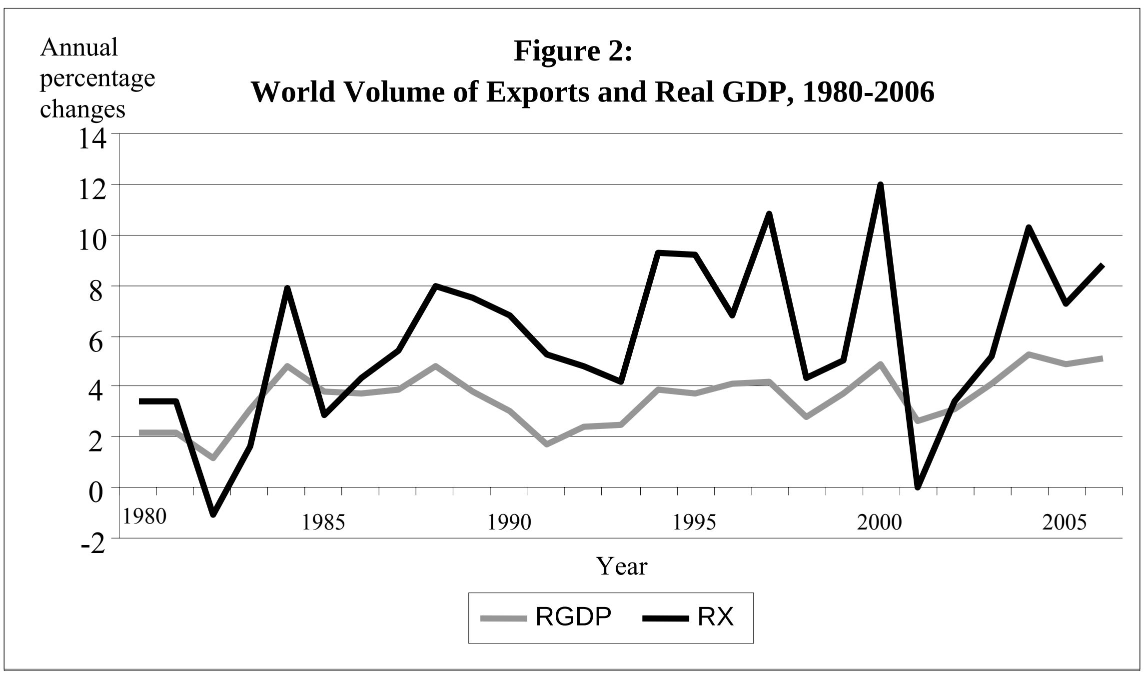
I. INTRODUCTION

International economics and international business have common interests but somewhat different research agendas. The former emphasizes cross-border trade and capital flows, whereas the latter looks predominantly at foreign direct investment. Part of this difference results from the emphasis that scholars in international business place on the study of the multinational firm and part is due to intellectual specialization. It is worth recalling that the yearly flows of international trade are a large multiple of the yearly flows of foreign direct investment, while the stock of foreign direct investment has only recently approached annual trade flows (see Figure 1). Furthermore, total real exports have grown faster, on average, than the world real GDP since the mid-1980s (see Figure 2). Finally, it is widely believed that exports are an engine of economic growth; see Krueger (2006). For all these reasons, international trade economists spend a great deal of time and resources understanding and explaining trade flows.

[Insert Figures 1 and 2 here]



Sources: International Monetary Fund (2006) and UNCTAD, http://stats.unctad.org/fdi/



Source: International Monetary Fund (2006).

With this brief background, I can state the objectives and outline of the chapter. The objective is to explain trade flows in terms of the gravity equation (GE). The reason for focusing on GE is two fold. The first is that GE, unlike other frameworks, has had great empirical success in explaining bilateral trade flows. For a long time, however, GE was a child without a father in the sense that it was thought to have no theoretical support. Since the late 1970s, this state of affairs has changed radically. Now, the gravity equation has strong theoretical support and can be derived from a variety of models of international trade. The second is that GE can be used to sort out alternative hypotheses of international trade.

In its simplest form, the gravity equation (GE) explains flows of a good between pairs of countries in terms of the countries' incomes, distance and a host of idiosyncratic factors--such as common border, common language, and common money-- that enhance or reduce bilateral trade flows:

$$(1) \qquad M_{ijk} = A_{0k} Y_{ik}^{\alpha_{1k}} Y_{jk}^{\alpha_{2k}} d_{ij}^{\alpha_{3k}} U_{ijk},$$

where M_{ijk} denotes that the k good is exported by country i and imported by country j, Y_{ik} and Y_{jk} are expenditures on the k good by the two countries, and d is distance; A and αs are coefficients, and U is a well behaved error term. The vector of idiosyncratic factors has been omitted in (1) because these factors are more control variables than theoretically derived variables. Aggregating over all k goods, the GE of a given product can be transformed into a GE of total exports of country i:

(2) $$M_{ijk} = A_0 Y_i^{\alpha_1} Y_j^{\alpha_2} d_{ij}^{\alpha_3} U_{ij},$$

where the k subscript has been suppressed and Y is the country's income (for example, nominal gross domestic product or GDP). The implications of GE –which we develop and discuss below-- are such that α_1 and α_2 are positive and in some instances equal to one and that α_3 is negative. Typically, equation (2) is specified in log linear form and estimated either with cross-section or panel data. In the latter case, a time subscript τ is added, except for the time-invariant physical distance:

(3) $$\ln(M_{ij\tau}) = \alpha_0 + \alpha_1 \ln(Y_{i\tau}) + \alpha_2 \ln(Y_{j\tau}) + \alpha_3 \ln(d_{ij}) + \alpha_4 F_{ij} + u_{ij\tau},$$

where ln stands for the natural logarithm, $\ln(A_0) = \alpha_0$ and $u_{ijt} = \ln(U_{ij\tau})$. The vector of idiosyncratic factors, F_{ij}, has also been added to equation (3). These factors are typically measured as dummy variables that acquire the value of one for the existence of the phenomenon and zero for its absence. The coefficients α_1 through α_3 are interpreted as elasticities or as percentage changes in bilateral trade for one percentage change in income and distance. The coefficient α_4 is positive if the factor is trade enhancing (e.g., common language) and negative if trade reducing (e.g., terrorism).

In the following section I will explore different models of international trade from which GE can be derived, ranging from models of complete specialization and identical consumers' preferences (Anderson 1979; Bergstrand 1985; Deardorff 1998) to models of product differentiation in a regime of monopolistic competition (Helpman 1987) to hybrid models of different factor proportions and product differentiation (Bergstrand

1989; Evenett and Keller 2002) to models of incomplete specialization and trading costs (Haveman and Hummels 2004).

II. TRADE THEORY AND THE GE

Complete specialization

Specialization is at the heart of trade theory; it is complete or deepest when each country specializes in the production of its own output and consumers purchase the output of each country according to identical and homothetic preferences. Furthermore, trade occurs without friction, meaning that it is not impeded either by transport costs, tariffs or tariff-equivalent border obstacles. This idealized set-up serves the purpose of creating a benchmark of maximum trade flows. Each country imports and consumes a share of the goods produced by all other countries, as well as a share of its own output. These shares are the same for all countries. Consider, for example, two countries, country 1 and country 2, producing differentiated products by country of origin. Country 1 will export its own good to country 2 in the amount of $M_{12} = b_1Y_2$, where b_1 = marginal propensity to import good 1 in country 2. Country 1 will also sell b_1Y_1 amount of the good it produces to domestic consumers. Note that the propensity to consume good 1 is the same across all consumers regardless of location. Income of country 1 is the sum of purchases by consumers located in country 1 and consumers located in country 2, i.e., $Y_1 = b_1Y_1 + b_1Y_2 = b_1Y_w$, where Y_w = world income = $Y_1 + Y_2$. Thus, identical and homothetic preferences imply that the propensity to import and consume good 1 is equal to country 1's share of world income. Replacing b_1 with Y_1/ Y_w, $M_{12} = Y_1Y_2/ Y_w$. This is the simple GE derived by Anderson (1979, p. 108):

(4) $M_{ij} = Y_i Y_j / Y_w$.

Bilateral trade flows respond positively to the incomes of the trading partner with a unit elasticity (one percent change in Y raises M by one percent) and negatively to world income. Referring to our empirical GE (3) and ignoring all other variables, complete specialization in trade implies that $\alpha_1 = \alpha_2 = 1$ and that the intercept of the regression can be interpreted as $-\ln(Y_w)$.

Introduce now trading costs ---a collection of costs that includes transaction, transport, and border-related costs-- such that exports are valued M_{ij} (f.o.b. prices) in country i but $M_{ij}\, t_{ij}$ (c.i.f. prices) in country j, where t_{ij} = (1 + trading costs per unit of exports). For the importing country, M_{ij} is now equal to $Y_i Y_j / t_{ij}\, Y_w$. Trading costs are not observable and the usual assumption is to proxy these costs by the distance separating the two countries, d_{ij}: more precisely, $1/t_{ij} = d_{ij}^{\alpha_3}$, with $\alpha_3 < 0$. The end result is that with trade frictions, bilateral trade flows fall by $1/t_{ij}$ or $d_{ij}^{\alpha_3}$ relative to the frictionless world of equation (4):

(5) $M_{ij} = Y_i Y_j / t_{ij}\, Y_w$.

Monopolistic competition and intra-industry trade

Complete specialization is the natural outcome of models of monopolistic competition and increasing returns to scale, which are at the core of the so-called New Trade Theory

(Helpman and Krugman 1985). Consumers like varieties and firms respond by differentiating their products by investing in a brand name. Separate markets develop for each of the differentiated products with the producer gaining some monopoly power and an ability to exploit economies of scale. As countries develop and mature, the demand for varieties increases and international trade tends to occur within the same industry. The industry has to be defined at a high level of disaggregation to avoid that different products may fall under the same classification. Trade overlap is measured typically with the Grubel-Lloyd index, which is the industry sum of twice the minimum of bilateral import values for each industry as a proportion of bilateral imports. The index is comprised between a minimum of zero for no intra-industry and a maximum of one fro complete intra-industry trade. A single-country measure of intra-industry trade activity is obtained by taking the weighted averages of bilateral indices. Data compiled by the OECD (2002, Table 2), at the fairly aggregated 2-digit SITC level, show that the average intra-industry trade of the 29 member countries exceeded 60 per cent of total manufacturing industry trade in the period 1996-2000.[1] Lionel Fontagné and Michael Freudenberg (1999) analyze 10,000 products at the highly disaggregated 8-digit category of the Combined Nomenclature of Eurostat for European countries for 1980, 1987, and 1994. Product similarity or overlap is judged in terms of differences in unit values. If export and import unit values differ by less than 15 per cent, the cross-border flows are defined as intra-industry trade. Based on this criterion, in 1994 intra-industry trade as a proportion of total intra-EU trade was close to 70 per cent for France, Germany, and Belgium-Luxembourg, 50 per cent for Italy and Spain and 14 per cent for Greece.

[1] There is a wide distribution, ranging from 77 per cent of the Czech Republic to 20 per

Furthermore, when these authors aggregate the 10,000 products into 14 industries, they find that intra-industry trade rises, over time, in all sectors except for agriculture and automobiles.

Helhanan Helpman (1987), drawing on his work with Paul Krugman (Helpman and Krugman 1985, ch. 8), develops a model that directly addresses intra-industry trade.[2] The key testable implication is that intra-industry trade responds positively, not only to the level of aggregate income, but also to the degree of income similarity among trading partners. More specifically, for a group of developed countries, such as those belonging to the OECD, Helpman develops and tests the following equation:

$$(6)\ X_A / Y_A = (Y_A / Y_w)[1 - \sum_{j\epsilon A}(b_{j,A})^2],$$

where X_A = trade inside the designated group A (e.g., OECD countries), Y_A = income of group A, and $b_{j,A}$ = income share of country j in group A. The expression in squared brackets measures the degree of income asymmetries in the designated group of countries. If one country were to be extremely big in relation to the others, $\sum_{j\epsilon A}(b_{j,\ A})^2$ would approach unity and the expression in the squared brackets would tend to zero. As countries in the group become less asymmetric in income, the expression in squared brackets rises. Thus, the more symmetric the group's countries are, the larger is the trade volume within the group. David Hummels and James Levinsohn (1995, pp. 804-5) test a slightly modified (6) applied to country pairs:

---

cent for Ireland.

[2] The GE from a model of product differentiation and monopolistic competition was also developed by Bergstrand (1989). In Bergstrand's, there is the added feature that bilateral

$$(7)\ \ln(M_{i+j}) = \alpha_1 \ln[Y_{i+j}\,(1 - b_i^2 - b_j^2)] + \ln\,(Y_{i+j}/Y_w),$$

where M_{i+j} is the volume of trade in the country pair and Y_{i+j} is the income in the two-country region. The first term on the right-hand side of (7) captures the impact of the country pair's income on bilateral trade. The income variable is corrected by the symmetry factor $(1 - b_i^2 - b_j^2)$. As countries become more similar the impact of income on bilateral flows rises. The second term of equation (7) captures the share of the pair's income in world income. The empirical results show that (7) works just as well for non-OECD as it does for OECD countries. Since intra-industry trade is small among developing countries, one must infer that there is more than differentiated goods in driving trade flows.

Hecksher-Ohlin

Equation (4) has also been derived from the perspective of the Hecksher-Ohlin (H-O) theory of comparative advantage (Deardorff 1998). This theory, as every undergraduate student of international economics and business knows, underscores the importance of a country's relative resources in determining its comparative advantage. In the H-O world, goods are homogeneous and perfect competition prevails. H-O predicts that a country will export those goods that require a relative intensive use of the endowed input. Capital-rich countries enjoy a lower cost of capital relative to wages and tend to export capital-intensive products; the reverse is true for labor-rich countries. There is also a

---

trade responds negatively to the size of the population of the two trading partners.

strong version of the H-O model, one in which input prices –i.e., cost of capital and wages—are equalized across countries. In this case, every country uses the same ratios of inputs and product prices do not differ. The amount of trade is not determinate. How is it then possible to determine bilateral trade flows predicted by the incomes of the two trading partners, as in our equation (4)? Deardorff arrives at a solution by invoking the principle of random separation of imports and exports. In his own words (pp. 12-3): "If consumers draw [randomly]…, then the law of large numbers will allow to predict quite accurately what their total choices will be by using expected values. In general, these expected values will be appropriate averages of the wide variety of outcomes that are in fact possible in the model."[3] An alternative route to arrive at (4), using the H-O framework, would consider complete specialization emanating from large differences in factor composition, a theme on which we will return later in the paper.

Incomplete specialization

So far, we have seen the gravity equation from complete specialization models. These models predict that a producer of a given good will supply all consumers or all countries. Consequently, we should note that the matrix of bilateral trade is full in the sense that an exporter will satisfy all importers. What is the evidence on this score? Ideally, as Haveman and Hummels point out (2004, p. 213), we would want to have data showing the complete range of varieties produced and cases when a country produces a good but

[3] Even though production costs are the same across countries, capital-rich countries will produce a disproportionate share of capital-intensive goods, and the opposite for labor-intensive countries. Thus, although factor prices are equalized, with consumers having identical and homothetic preferences, the main H-O prediction holds (Helpman 1989, p. 124).

fails to export it. This information is unfortunately not available. The procedure adopted by Haveman and Hummels is to define a good at a 4-digt SITC category, compute the number of exporters of good k for each importer and then divide this number by the total number of exporters of good k. Under complete specialization, this number should be equal to one. Instead, these authors find that 27 per cent of the sample has zero values and 58 per cent of importers buy from fewer than 10 per cent of available exporters.[4] In sum, the foundation upon which complete specialization is based may be a bit shaky. The next step involves deriving the gravity equation from the alternative perspective of incomplete specialization, an environment in which consumers buy a sub-set of available varieties or there are multiple suppliers of homogeneous goods.

Simon Evenett and Wolfgang Keller (2002) derive two GE-like testable models from incomplete specialization. In the first one, the setting is H-O but restricted to two goods, two factors, and two countries. The restriction is essential because in a more plausible many-country environment bilateral trade flows with multiple suppliers (in a frictionless world) would be indeterminate. The two countries have different capital to labor ratios and export different goods. Their bilateral exports will depend not only on income, as in equation (4), but also on the share of the two goods in production (see equation 3 in Evenett and Keller):

$$(8)\ M_{ij} = (b_i - b_j)/(Y_i Y_j / Y_w),$$

where b defines the share of one of the two goods in production. Say b defines the share of the capital-intensive good and country i is relatively rich in capital, then $b_i > b_j$ and exports of the capital-intensive good go from i to j. For the labor-intensive good, the

[4] The exercise is done for the year 1990, with 75,774 observations covering 173 countries

shares are $(1 - b_i)$ and $(1 - b_j)$, respectively; exports of the labor-intensive good flow from j to i. As the factor mix in the two countries converge, b_i and b_j become more equal and bilateral trade peters out. In the limit, in the absence of differences in factor endowments, trade disappears altogether. Note that (8) implies a lower elasticity of trade with respect to income than equation (4) derived from complete specialization.

The alternative way to generate a GE from incomplete specialization is to assume that a country produces a homogeneous good under constant returns to scale, as envisioned in the H-O world, as well as a differentiated good under increasing returns to scale, as envisioned by the New Trade Theory. In this case, the testable implication is (see Evenett and Keller, equation 2):

$$(9)\ M_{ij} = (1 - b_i)/(Y_i Y_j / Y_w),$$

where b_i is the share of the homogeneous good in country i. As b_i approaches zero, equation (9) converges to equation (4).

It is useful to compare equation (4), derived under complete specialization and trading costs, with equation (8), derived under a restrictive H-O environment, and with a specialized H-O environment, and with equation (9), derived under a mixed environment of differentiated good and H-O. Income predicts the highest trade volume in the complete specialization case, followed by the mixed case, and lastly by the restricted H-O model; see Evenett and Keller (2002, pp. 286-7). We shall return to this point in the next section.

---

and 438 4-digit sectors.

III. THE EMPIRICAL GE

The GE has been very successful in explaining actual trade patterns; in fact, it is considered to be state of the art for the determination of bilateral trade (Leamer and Levinsohn 1995, p. 1384; Feenstra et al. 2001, p. 431). There is a voluminous literature on the empirical GE going back to the early 1960s, far too big to be reviewed within the space of a chapter. The approach I shall follow is to select some themes that are germane to the theory presented above.

GE and alternative trade theories

Among the various uses of the GE, one of the most promising is to employ it so as to discriminate among alternative theories of international trade. This is what Feenstra et al (2001) do. These authors fit a GE of the form of equation (3) to 4-digit SITC level trade data from Statistics Canada World Trade Analyzer (WTA). The disaggregated data are grouped in three categories following Rauch's (1999) methodology. The first group consists of homogeneous goods traded in exchanges and whose prices are very transparent; the second group consists of reference-price goods whose prices can be obtained in industry publications; and the third group consists of 'differentiated' goods with unquoted prices. In other words, products are ranked by degree of homogeneity and price transparency. Both homogeneous and differentiated goods industries have barriers to entry. Entry barriers in the homogeneous good industries tend to be more in the form of large fixed costs of capital outlays, where in the differentiated goods industry are in the form of fixed costs to develop brand names. If entry barriers are uncorrelated across homogeneous and differentiated goods industries, the null hypothesis (H_o) is that α_1 and

α_2 of equation (3) are not significantly different from each other. The authors consider two alternative hypotheses to H_o: one (H_1) in which entry barriers are higher in the differentiated goods industry than in the homogeneous goods industry and the other (H_2) where entry barriers are higher in the homogenous goods industry. Under H_1, $\alpha_1 < \alpha_2$; the higher income elasticity of the importer can be intuitively justified by what is called a 'reverse' home market effect. Under H_2, $\alpha_1 > \alpha_2$; the higher income elasticity of the exporter can be intuitively justified by what is called a home market effect. The evidence presented by the authors rejects the statistical equality of the two alpha coefficients. The estimated α_1 (exporter's income) rises as one moves from homogeneous to reference-priced goods to differentiated goods. In the homogeneous category, α_1 is below α_2, whereas the opposite is true for differentiated goods. In sum, the evidence is strongest for H_2, that is the environment where barriers to entry are highest, such as in sectors like mining and steel.

Complete vs. incomplete specialization

There is some evidence that the data fit better models of incomplete specialization, as implied by our equation (8) and (9), than models of complete specialization, as implied by equation (4). To test this proposition, Evenett and Keller (2002) apply the Grubel-Lloyd index to construct an index of intra-industry trade using 1985 trade data at the 4-digit SITC level from 58 (primarily industrial) countries. As already noted, the intra-industry index is bound between zero (total absence of intra-industry trade) and one (trade takes only the form of intra-industry varieties). In practice, the index tends to be low. In the authors' sample, the distribution of the index is skewed towards zero, with 78 per cent of the sampled country pairs lying below the value of 0.05, which is taken by the authors as the demarcation line between homogeneous and differentiated production. Equation (4) is applied to data whose intra-industry index falls below 0.05. The unit coefficient on the incomes of the two trading partners is rejected by the data. We recall that equation (4) could have been generated from either differentiated varieties produced under increasing returns to scale or by homogenous produced in an H-O world with large differences in factor endowments. Empirically, the alternative of incomplete specialization fares much better. Equations (8) and (9) are tested with data whose intra-industry trade index exceeds 0.05. For equation (9), the coefficients of the income variables should be positive (but below unity) and rising as the Grubel-Lloyd index rises. The findings do not reject these patterns. Equation (8) is tested for different classes of factor intensities –specifically different values of capital to worker ratios and a Grubel-Lloyd index below 0.05--. Here as well the findings are consistent with equation (8). In sum, models of complete specialization overpredict bilateral trade flows and are rejected in favor of models of incomplete specialization.

National borders and multilateral resistance

National borders are a discontinuity of distance and an impediment to international trade. Costs take a jump at the border. First, there are transaction costs due to customs clearance and formalities. Furthermore, the border is a delimiter of differences in legal systems and practices, languages, networks, competitive policies, and monetary regimes. Finally, national authorities use the border to discriminate against foreign producers by applying tariffs or tariff-equivalent restrictions. Border frictions are more difficult to quantify than distance-related frictions.

The economic size of the national border is at center stage in McCallum (1995) who fits a modified form of equation (3) to 1988 exports and imports among ten Canadian provinces and 30 U.S. states. In addition to income and distance, McCallum's GE codes a dummy variable equal to one for inter-provincial trade and zero for province-to-state trade. The point estimate of the dummy variable, the size of the border effect, is approximately 3 and statistically significant under a variety of tests. Since the exponent of the coefficient of the dummy variable is approximately 20, McCallum's findings imply that inter-provincial trade (i.e., trade within Canada) is approximately twenty times larger than trade between provinces and states (i.e., trade between Canada and the United States).

Anderson and van Wincoop (2003) have criticized McCallum's results of very thick borders for ignoring the asymmetric impact on trade of barriers between small and large economies and multilateral protection levels. On the first point, these authors re-estimated the gravity equation, using McCallum's exact specification and data, alternatively from the viewpoint of Canada and the United States, and found that the border from the Canadian viewpoint is ten times as wide as from the viewpoint of the United States. Since Canada's economy is approximately one-tenth of the United States', the level of protection imbedded in a border is a positive function of the size of the economy. Helliwell (1996) has confirmed these findings with data for the province of Quebec, the author desiring to underscore, among other things, what Quebec would lose from a possible separation from Canada.

At a more fundamental level, Anderson and van Wincoop criticize those specifications of GE that ignore the interaction between bilateral and multilateral trading costs. Their basic contention is that bilateral trade flows depend on what goes on between a given country pair and the rest of the world; in other words, bilateral trade flows are determined by a general equilibrium framework. When multilateral trading costs rise relative to bilateral costs, trade flows rise between the country pair i and j; and vice versa. The authors arrive at a testable equation that resembles equation (5) above:

(10) $$M_{ij} = (Y_i Y_j / Y_w)(t_{ij}/P_i P_j)^{(1-\sigma)}.$$

There are two differences between (10) and (5). The first concerns trading costs. We have seen that Anderson and van Wincoop criticize McCallum's results on the ground that they ignore multilateral aspects of such costs. In (10), t_{ij} appears now in the numerator and it is deflated by price indexes of the trading partners, P_i and P_j, which in turn depend on all t_{ij} pairs, countries' income shares and countries' price levels. Thus, bilateral trading costs are divided by what Anderson and van Wincoop call 'multilateral resistance' factors. The second change concerns the exponent of the ratio of bilateral to multilateral resistance factors. The exponent is not one but (1-σ). The reason for this change comes from the fact that the hypothesized utility has a constant-elasticity of substitution, σ, between goods. This elasticity has been estimated to exceed one; the authors assume it to be five. In sum, an increase in multilateral resistance relative to bilateral resistance raises bilateral trade. The rest of (10) matches (5). This is because (10), like (5), was derived under the assumption of complete specialization.

Define $t_{ij} = \delta_{ij} d_{ij}^{\rho}$, where d_{ij} is, as before, distance between the country pair, and $\delta_{ij} = 1$ if the two trading regions are located inside a given border or $\delta_{ij} = 1$ plus the tariff rate and if the two regions are located on opposite sides of a border. Substituting the definition of t_{ij} in (10), we obtain

(11) $$\ln(M_{ij}/Y_i Y_j) = -\ln(Y_w) + (1-\sigma)\rho\ln(d_{ij}) + (1-\sigma)\ln(\delta_{ij}) - (1-\sigma)\ln(P_i) - (1-\sigma)\ln(P_j).$$

P_i and P_j are a function of bilateral distance, border, and other unobserved variables that influence trading costs. The Ps have to be estimated for all the countries; for that Anderson and van Wincoop use nonlinear least squares to minimize the sum of squared errors. A simpler, but less efficient, alternative is to use country-specific dummies; more on this below. With all these adjustments, the authors obtain that for the bilateral trade between Canada and the United States the tariff-equivalent border rate is 51 per cent.[5]

In sum, Canada and the United States have intense and relatively open trade relations; yet, their border represents a considerable obstacle to further integration. To have a more complete picture of the border effect, we need to extend the work to other countries, both in the North and the South.

<u>North and South trade</u>

Our last empirical topic deals with the application of GE to trade flows among developed countries (North), developing countries (South), and between the two sub-groups (North-South).

5 See the authors' table 2. The estimate of $(1-\sigma)\ln(\delta_{USCan})$ is -1.65. Given that σ is set to 5,

For this purpose, I shall use the WTA by Statistics of Canada, which consists of 215,500 annual observations on bilateral imports, in U.S. dollars, covering 143 countries over the period 1980 to 2003. Details of the data set can be obtained by consulting the Indiana University CIBER Website (http://www.kelley.iu.edu/ciber/research.cfm). The testable equation is a modified version of (3)

$$\textbf{(12)}\quad \ln(M_{ijt}) = \alpha_0 + \alpha_1 \ln(Y_i Y_j)_t + \alpha_2 \ln(Y_i Y_j / N_i N_j)_t + \alpha_3 \ln(d_{ij}) + \alpha_4 F_{ij} + \alpha_5 \ln P_{it} + \alpha_6 \ln P_{jt} + u_{ijt}.$$

Specification (12), unlike (3), imposes the restriction that the elasticity of trade flows with respect to income is the same for exporting and importing countries. In addition, (12) includes per-capita income and time-varying multilateral resistance factors P_{it} and P_{jt}. Bergstrand (1989) shows the relevance of per-capita income, which proxies for factor intensities in the GE. The multilateral trade factors were discussed in connection with the empirical work on the border effect. Unlike Anderson and Van Wincoop, I will account for these factors by the simpler procedure of using time-varying country-specific dummies (that is, country dummies interacting with years). A more comprehensive treatment of the econometric issues underlying the GE estimation with panel data can be found in Fratianni and Oh (2007).

Vector F includes affinity variables that are trade-enhancing. These fall into three categories: geographic affinity (common land border), cultural affinity (common

---

$\ln(\delta_{USCan}) = 0.4125$ and $\delta_{USCan} = 1.51$ (note that exp(0.4125) = 1.51).

language, common colonizer, and colonial relationship), and institutional affinity (RTA membership and common currency). RTAs work like clubs and give members privileged access to a geographic area. There is a big literature on whether RTAs, on balance, create trade or divert trade against outsiders; see Fratianni and Oh (2007). Regionalism is defined in terms of eleven separate RTAs, covering 40 per cent of world trade; for a list of the RTAs, see Table 1 below. Since the RTAs have the potential to divert trade, (12) includes also an inter-regional dummy that is equal to one when the trading partners belong to different RTAs; otherwise it is zero. Trade diversion implies a negative coefficient.

Measurement and sources of the variables of equation (12) are shown in the Appendix. Table 1 presents summary statistics for the entire sample of countries as well as for the North, South, and North-South sub-groups. North is defined as the set of countries that are members of the OECD; South is defined as the non-OECD countries. The average bilateral import is \$5.2 million, but in the North the average rises to \$458.7 million, whereas in the South falls to \$1.6 million. Just as stark are the differences in income. The average multiplicative GDP (Y_iY_j) in the North is 188 times the average for the South; the average multiplicative per-capita income in the North is 95 times the average for the South. Average distance is 4,589 miles, with a range spanning from 55 miles (Bahrain and Qatar) to 12,351 miles (Guyana and Indonesia). Country-pair observations with RTA membership represent 21.3 per cent of the sample in the North but only 2.5 per cent in the South; those with a common land border represent 7.4 per cent in the North and 4.5 per cent in the South; those with a common currency 3.4 percent in the North and 1.4 per cent in the South; those with a shared language 12 per cent in the North and 28 per cent in the South; those with a common colonizer are entirely located in the South; and those with a colonial relationship are much more frequent in the North than in the South.

[Insert Table 1 here]

Estimates of equation (12) are presented in Table 2. Income and distance are powerful forces of bilateral trade and appear to be stable across different groups of countries. The elasticity of bilateral imports with respect to GDP is between 1.10 and 1.20, the differences between North and South being quite minor. The elasticity of bilateral imports with respect to distance is -.99 for the North and -1.11 for the South (and the difference is statistically significant at the 1 per cent level), suggesting that trading costs are higher for developing countries than developed countries. To have an appreciation of the quantitative importance of distance, consider that the average of the log of distance is about 8.2 and that the average of the log of bilateral imports is 8.6. Since the distance elasticity is around unity, distance alone, on average, "destroys" almost the entire value of bilateral flows. Distance, we recall, captures more than mere transportation costs. The consensus is that the bulk of trading costs are due to trade-reducing factors such as differences in legal systems, administrative practices, market structures, networks, languages and monetary regimes; see Grossman (1998, pp. 30-31).

Membership to an RTA raises bilateral trade flows much more for the South and the North-South than for the North. On the other hand, the relatively low frequency of RTAs in the South may weaken the reliability of the estimates for South-South and North-South; for a more complete discussion of RTAs on trade flows, see Fratianni and Oh (2007). The inter-regional dummy is positive and statistically significant, suggesting that RTA membership has not hampered trade between countries that belong to two different trade clubs.

The relationship between bilateral trade and a common currency appears to be unstable: it is strongly positive in the South, but statistically insignificant in the North. Again, as it is true for the RTAs, the reliability of these estimates may reflect the fact that frequency of currency unions in the South is very low both in an absolute sense and in relation to the frequency in the North. At this point, it is best to remain cautious on the quantitative importance of a common currency on trade. The initial estimates of Rose (2000) that countries with a common currency trade three times as much as countries with different currencies (and fluctuating exchange rates) has been met with some skepticism; see, for example, the comments to Rose by Persson (2001).

Geographical affinity, proxied by a shared land border, is trade enhancing for the South but not for the North. The implicit assumption that a shared border leads to more trade is based on the presumption that neighboring countries have friendly relations and tend to cooperate more than distant countries. The alternative that close countries tend to be unfriendly and protect the home market more than distant countries cannot be dismissed, certainly not through reading of history. Cultural affinity, proxied by a common language, common colonizer and a shared colonial relationship expand trade across all sub-groups. Common language is more trade enhancing in the North than the South, while the opposite holds for colonial ties.

In sum, the results of equation (12) should be judged as a success for the explanatory power of the gravity equation.

[Insert Tables 2]

IV. CONCLUSIONS

The objective of this chapter was to provide a survey, albeit selective given space limitations, of the gravity equation in international trade. This equation started in the Sixties as a purely empirical proposition to explain bilateral trade flows, without little or no theoretical underpinnings. At the end of the Seventies, the gravity equation was rejuvenated and legitimized by a series of theoretical articles that demonstrated that the basic GE form was consistent with various models of trade flows. Empirical applications of GE expanded to cover a variety of issues, such as the impact of regional trade agreements, national borders and currency unions on trade, as well as the use of the equation to sort out the relative merit of alternative trade theories. A new wave of studies is now concentrating on the general equilibrium properties of the GE and finer econometrics points. The renewed interest of the academic profession in the development of the GE is undoubtedly driven by the equation's empirical success.

Table 1 Descriptive Statistics

Variable	Mean for Entire Sample	S.D. for Entire Sample	Mean for North-North	Mean for South-South	Mean for North-South
Number of observation	215,500	215,500	13,287	95,278	106,935
Log of bilateral imports	8.5643	3.2514	13.0362	7.3605	9.0813
Log of GDP	48.4699	2.7089	52.3583	47.1183	49.1911
Log of per capita GDP	15.9096	2.0875	19.1618	14.6053	16.6676
Log of distance	8.1953	0.7794	7.6219	8.1175	8.3349
Common RTA[a]	0.0258	0.1587	0.2131	0.0245	0.0038
Inter-regional	0.1304	0.3368	0.1774	0.0812	0.1685
Common land border	0.0268	0.1616	0.0741	0.0456	0.0042
Common currency[b]	0.0091	0.0948	0.0339	0.0137	0.0019
Common language	0.2140	0.4101	0.1199	0.2844	0.1629
Common colonizer	0.0804	0.2719	NA	0.1818	NA
Colonial relationship	0.0232	0.1505	0.0351	0.0021	0.0405

Notes: [a] The eleven RTAs are: the European Union, the North American Free Trade Association, the Association of South East Asian Nations, the Southern Common Market, the Caribbean Community and Common Market, the Andean Community of Nations, the Australia-New Zealand Closer Economic Relations Trade Agreement, the Central American Common Market, the Papua New Guinea-Australia Trade and Commercial Relations Agreement, the South Pacific Region Trade and Economic Cooperation Agreement, and the United States-Israel Free Trade Agreement; for more details see Fratianni and Oh (2007). [b] The list of monetary unions encompass the following areas: the U.S. dollar, the East Caribbean dollar, the Australian dollar, the Rihal, the euro, the CFA, the Franc, the Indian Rupee, and the Rand; for more details see Fratianni (2006).

Table 2 Estimates of equation (12). Dependent variable: log of bilateral imports. Sample period: 1980-2003.

Independent var.	Entire Sample	North-North	South-South	North-South
Intercept	-32.3019***	-30.9683***	-31.8036***	-35.9988***
	(0.1128)	(0.3276)	(0.2110)	(0.1833)
Log of GDP	0.9770***	0.8602***	0.9562***	1.0433***
	(0.0026)	(0.0056)	(0.0047)	(0.0034)
Log of per capita GDP	0.1342***	0.3323***	0.1923***	0.1158***
	(0.0034)	(0.0125)	(0.0060)	(0.0055)
Log of distance	-1.0837***	-0.9863***	-1.1065***	-1.0059***
	(0.0068)	(0.0124)	(0.0108)	(0.0111)
Common RTA	1.0562***	0.1863***	1.9642***	3.7129***
	(0.0292)	(0.0269)	(0.0518)	(0.0900)
Inter-regional	0.3440***	0.0958***	0.0831***	0.3739***
	(0.0154)	(0.0291)	(0.0320)	(0.0182)
Common currency	0.2581***	0.0058	0.5906***	0.1864
	(0.0473)	(0.0556)	(0.0696)	(0.1238)
Common border	0.4971***	-0.0964**	0.6595***	0.9029***
	(0.0289)	(0.0377)	(0.0395)	(0.0828)
Common language	0.3933***	0.4123***	0.2433***	0.5385***
	(0.0121)	(0.0299)	(0.0194)	(0.0169)
Common colonizer	0.7185***	NA	0.6592***	NA
	(0.0184)		(0.0237)	
Colonial relationship	1.2167***	0.4958***	1.0247***	1.1521***
	(0.0300)	(0.0469)	(0.1568)	(0.0300)
Obs.	215,500	13,287	95,278	106,935
R^2	0.6557	0.8543	0.4686	0.6910

Note: Equation (2) has been estimated with time-varying importing country fixed effects, which are not reported. Superscripts *,**, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively. Dependent variables and GDP are in nominal dollar values.

REFERENCES

Anderson, James E. 1979. A Theoretical Foundation for the Gravity Equation. *American Economic Review* 69: 106-16.

Anderson, James E. and van Wincoop, Eric. 2003. Gravity with Gravitas: A Solution to the Border Problem. *American Economic Review* 93 (March: 170-192.

Bergstrand, Jeffrey H.1985. The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics* 67(3): 474-481.

Bergstrand, Jeffrey H.1989. The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade. *The Review of Economics and Statistics* 71(1): 143-153.

Deardorff, Alan V. 1998. Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World? In Jeffrey A. Frankel (ed.), *The Regionalization of the World Economy*. Chicago: University of Chicago Press (for NBER).

Evenett, Simon J. and Wolfgang Keller. 2002. On Theories Explaining the Success of the Gravity Equation. *The Journal of Political Economy* 110(2): 281-316.

Feenstra, Robert C., Markusen James R. and Andrew K. Rose. 2001. Using the Gravity Equation to Differentiate among Alternative Theories of Trade. *The Canadian Journal of Economics* 34(2): 430-447.

Fontagné, Lionel and Michael Freudenberg. 1999. Endogenous Symmetry of Shocks in a Monetary Union. *Open Economies Review* 1999, 10: 263-287

Fratianni, Michele. 2006. Borders and Integration. In Michele Fratianni (ed.), *Regional Economic Integration*. Amsterdam: Elsevier JAI.

Fratianni, Michele and Chang Hoon Oh. 2007. Do Expanding RTAs Behave like Trade Building Blocs?, Indiana University Business Economics and Public Policy Working Paper 2007-13.

Grossman, Gene M. 1998. Comment to Alan Deardorff's Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World? In Jeffrey A. Frankel (ed.), *The Regionalization of the World Economy*. Chicago: University of Chicago Press (for NBER).

Haveman, Jon and David Hummels. 2004. Alternative Hypotheses and the Volume of Trade; the Gravity Equation and the Extent of Specialization. *The Canadian Journal of Economics* 37(1): 199-218.

Helliwell John F. 1996. Do National Borders Matter for Quebec's Trade? *Canadian Journal of Economics 29*(3): 507-522.

Helpman, Elhanan. 1987. Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries. *Journal of the Japanese and International Economies* 1:62-81.

_______________. 1989. The Structure of Foreign Trade. *The Journal of Economic Perspectives* 13(2): 121-144.

Helpman, Elhanan and Krugman, Paul. 1985. *Market Structure and Foreign Trade*. Cambridge, MA: MIT Press.

Hummels, David and James Levinsohn. 1995. Monopolistic Competition and International Trade: Reconsidering the Evidence. *The Quarterly Journal of Economics* 110, 3: 799-836.

International Monetary Fund. 2006. World Economic Outlook Database for September 14, 2006; available at http://www.imf.org/external/pubs/ft/weo/2006/02/data/download.aspx)

Krueger, Anne O. 2006. A Remarkable Prospect: Opportunities and Challenges for the Modern Global Economy," Lecture at the Claremont McKenna College Claremont, California (May 2), available at http://www.imf.org/external/np/speeches/2006/050206.htm

Leamer, Edward E. and James Levinsohn. 1995. International Trade Theory: The Evidence. In Gene M. Grossman and Kenneth Rogoff (eds.), *Handbook of International Economics*. Amsterdam: Elsevier.

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

OECD. 2002. Intra-industry and Intra-firm trade and the Internationalisation of Production, *Economic Outlook* 71: 159-170.

Persson, Thorsten. 2001. Currency Unions and Trade: How Large is the Treatment Effect? *Economic Policy*, 33: 435-448.

Rauch, James R. 1999. Networks versus markets in international trade, *Journal of International Economics* 48:7-27.

Rose, Andrew K. 2000. One Money, One Market: The Effects of Common Currency on Trade. *Economic Policy 30,* (April) 9-45.

United Nations Conference on Trade and Developemnt (UNCTAD). Foreign Direct Investment, FDI Flows and Stocks, available at http://stats.unctad.org/fdi/.

Appendix 1 Data description

Variables	Descriptions	Data Sources	Units
Bilateral Imports	Log of nominal bilateral imports	World Trade Analyzer[a]	Log of 1000 US dollar
Log of nominal GDP	Log of the product of nominal GDPs.	World Development Indicator[b]	Log of dollar
Log of nominal per capita GDP	Log of the product of nominal per capita GDPs.	World Development Indicator[b]	Log of dollar
Log of Distance	Log of distance between trading partners	World Factbook[c]	Log of mile
Common Border	If two countries share a common border, Common Border = 1, otherwise 0.	World Factbook[c]	Dummy variable.
Common language	If two countries share same main language, Common language = 1, otherwise 0.	World Factbook[c]	Dummy variable.
Common Colonizer	If two countries had same colonizer, Common Colonizer = 1, otherwise 0.	World Factbook[c]	Dummy variable.
Colonial Relationship	If two countries were involved in a colonial relationship with each other, Colonial Relationship = 1, otherwise 0.	World Factbook[c]	Dummy variable.
Common Currency	If two countries share the same currency or a unit exchange rate, Common Currency = 1, otherwise 0.	IMF[e]	Dummy variable.
Same-RTA (11 RTAs)	If two countries belong to the same RTA in the year of observation, Same-RTA = 1, otherwise 0.	WTO[d]	Dummy variable.
Inter-regional	If exporting and importing countries belongs to different RTAs, inter-regional = 1, otherwise 0.	WTO[d]	Dummy variable.

Notes:

[a] "World Trade Analyzer" (WTA) has been assembled and managed by Statistics Canada. Information of the data is available at http://www.statcan.ca/english/ads/trade/world.htm

[b] The source for nominal GDP is World Bank's "World Development Indicators". When data are unavailable from World Bank, missing observations are filled from the "Penn World Table" and IMF's "International Financial Statistics".

[c] "World Factbook", CIA; http://www.cia.gov/coa/publication/factbook

[d] The data available at http://www.wto.org/english/tratop_e/region_e.htm

[e] The basic source for currency unions is the IMF's "Schedule of Par Values" and issues of the IMF's "Annual Report on Exchange Rate Arrangements and Exchange Restrictions". Data are supplemented by the yearly "Statesman's Year Book".