The Gravity Equation in International Economics and International Business Research: A Note

By

Michele U. Fratianni*, Francesco Marchionne** and Chang Hoon Oh***

Abstract

This note discusses methodological issues and practical concerns for international economists and international business scholars who apply the gravity equation in their research. The most important message of the note is that this equation should correct for multilateral resistance factors. We propose a relatively low-cost specification and estimation to implement such correction, which is robust in the presence of various endogeneity effects and non-stationary variables. In the presence of zero-values in the dataset, however, the multilateral specification is best estimated with a Poisson maximum likelihood.

Keyword: gravity equation; international trade; foreign direct investment; methodology.

JEL Classification: C1, F1, F2

*Indiana University, Kelley School of Business, Bloomington, IN 47405 USA and Department of Economics, Università Politecnica delle Marche, Ancona, Italy, Email: fratiann@indiana.edu
**Department of Economics, Università Politecnica delle Marche, Ancona Italy, Email: f.marchionne@univpm.it
*** Faculty of Business, Brock University, St.Catharines , Ontario L2S3A1 Canada, Email: coh@brocku.ca.

INTRODUCTION

The gravity equation (GE) has been a strong work horse in explaining the determinants of bilateral trade and foreign direct investment (FDI) flows. The importance of GE extends beyond economics. For example, the *Journal of International Business Studies*, the leading review of international business, has published, from 1996 to 2009, 10 articles where GE has been explicitly employed to analyze FDI flows, trade flows, and both trade and FDI flows. The wider use of GE has been supported by theoretical and statistical contributions. This note intends to provide methodological best practices to the growing number of economists and international business scholars who are interested in the GE.

THEORY UNDERLYING THE GRAVITY EQUATION

There is a large literature showing that GE can be derived from a wide variety of frameworks of international trade, 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 & Keller, 2002) to models of incomplete specialization and trade costs (Haveman & Hummels, 2004).[1]

Under complete specialization, each country specializes in the production of its own output while consumers purchase the output of each country according to identical preferences. Furthermore, trade occurs without any friction of transport costs, tariffs or tariff-equivalent border obstacles. This idealized set-up serves the purpose of creating a benchmark of maximum trade flows given by the following equation:

(1) $$X_{ij} = Y_i Y_j / Y_W,$$

where X_{ij} denotes exports of country i to country j, $Y_i = i$'s income, $Y_j = j$'s income, and $Y_w =$ world income. Now, introduce trade costs --a collection of costs that includes transaction, transport, and border-related costs-- such that exports are valued at FOB prices in country i but at CIF prices in country j.[2] Equation (1) becomes:

(2) $$X_{ij} = Y_i Y_j / (t_{ij} Y_W).$$

where t_{ij} = one plus trade costs per unit of exports. Recall that trade costs are largely unobservable; the usual practice has been to proxy these costs by the geographical distance between two country-partners.

In models of monopolistic competition and increasing returns to scale (Helpman & 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 monopolistic 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. In this world, X_{ij} responds positively not only to level of aggregate income but also to the degree of income similarity between country–partners (Helpman, 1987).

Models of complete specialization predict that a producer of a given good will supply all consumers or all countries. In fact, the bilateral trade matrix has zeros and, consequently, the foundation upon which complete specialization is based may be a bit shaky (Haveman & Hummels, 2004). Within the alternative perspective of incomplete specialization, Evenett and Keller (2002) start from a restricted Hecksher-Ohlin set-up of two goods, two factors, and two countries, and arrive to the following GE:

(3) $$X_{ij} = (b_i - b_j)/(Y_i Y_j / Y_W),$$

where b defines the share of one of the two goods in total 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 shares are $(1 - b_i)$ and $(1 - b_j)$, respectively; exports of the labor-intensive good flow from j to i. As the factors converge in the two countries, b_i and b_j become more equal and bilateral trade peters out.

A recent important advancement in the GE has been proposed by Anderson and van Wincoop (2003), who argue that bilateral trade flows depend also on what goes on between a given country-pair and the rest of the world; in other words, bilateral trade flows are determined in a general equilibrium framework. These authors derive the following GE:

(4)
$$X_{ij} = Y_i Y_j / Y_W (t_{ij} / P_i P_j)^{1-\sigma},$$

where the new variable P denotes multilateral trade costs and the coefficient σ is the elasticity of substitution coefficient. For $\sigma > 1$, bilateral trade flows rise (fall) if multilateral trade costs rise (fall) relative to bilateral trade costs. P_i and P_j are a function of all t_{ij} pairs, countries' income shares, countries' price levels, and factors unobservable to researchers. Just as importantly, they are jointly determined and their omission creates a bias in the estimated coefficients. We will return to this critical issue later.

International business scholars have also used the GE extensively on their favorite topic of the determinants of FDI flows, see, for example, Hejazi and Pauly (2005), Buckley et al. (2007), Clougherty and Grajek (2008), Li and Vashchilko (2009), and others. Yet, as Li and Vashchilko (2009: 8) state, "the gravity model of FDI does not have as strong a theoretical foundation as in the case of trade…" On the other hand, the state of the art is in rapid evolution. We briefly review three articles. The first is by Brainard (1997) who seeks to explain the equilibrium outcome that foreign markets can be served either by exports or by sales of foreign affiliates of horizontally-expanded MNCs. In equilibrium, the share of total sales in the foreign market accounted for by exports as opposed to affiliates' sales is greater when transport costs and trade barriers are lower and the fixed cost of setting up a foreign plant is higher. Brainard tests, with some degree of success, a GE of export shares being a negative function of freight cost, tariffs, home taxes, and FDI openness, and a positive function of the differential per-worker GDP (capturing factor-proportion differences), and trade openness in the host country.

The second is by Egger and Pfaffermayr (2004) who bring to center stage the nature of the foreign affiliate. The theoretically expected effect of distance --proxying for differences in cultural, economic and political environments-- on exports and outward FDI depend not only on the importance of distance for FDI relative to exports but also on whether the MNC expands horizontally (produce and sell in the host country) or vertically (produce in the host country and sell in the home country, or *vice versa*, through intra-firm trade).[3] The comparative statics of the authors' three-factor model yield predictions for income, size similarity, relative capital endowment, relative human capital

endowment, relative labor endowment, and distance. Distance acts negatively on trade and positively on horizontal and vertical MNCs when it has a low impact on new plants relative to exports. When this relative impact of distance is instead high, the signs of the distance coefficients on the GE for exports and outward FDI are reversed in the case of horizontal MNCs, whereas they are both positive for exports and outward FDI in the case of vertical MNCs.

The third is by Kleinert and Toubal (2010) who demonstrate that the GE applied to affiliates' sales can be derived from three different models. The first is an extension of Brainard's (1997) and leads to a GE in which affiliates' sales are a positive function of Y_i, Y_j and P_j and a negative function of distance. Both income elasticities are constrained to unity just like in the trade GE. The second features heterogeneous firms and fixed costs of setting a foreign plant that increase with distance. This set-up leads to a GE of the same form of the first model. In both first and second model the MNC operates horizontally. The third model is based on factor-proportion and vertical MNCs. The derived GE in log-linear form includes, in addition to Y_i, Y_j and distance, the sum of the two incomes and relative factor endowments. Affiliates' sales respond with a unit elasticity to the sum of the two incomes, positively to Y_j and relative factor endowments, and negatively to Y_i and distance.

In sum, the trade GE has a longer history and deeper theoretical underpinnings than the FDI GE, but recent developments have reduced the gap between the two. Clearly, more work is needed in this area.

ESTIMATION ISSUES

We consider four key econometric issues that the literature has raised with respect to the GE estimation. We organize our discussion around the following testable equation for panel studies:

(5) $\ln(X_{ijt}) = \alpha_0 + \alpha_1 \ln(Y_{it}) + \alpha_2 \ln(Y_{jt}) + \alpha_3 \ln(d_{ij}) + \alpha_4 CC_{ijt} + \alpha_5 \ln P_i + \alpha_6 \ln P_j + \alpha_t + u_{ijt}.$

The new terms are as follows: d is the cost-increasing distance; CC is a vector of dummy variables capturing cost-reducing affinities shared by country-pairs; α_t is a time effect common to all country-pairs; and $u_{ijt} = x_i + x_j + \mu_{ij} + \varepsilon_{ijt}$, where x_i and x_j are respectively exporter and importer time-invariant fixed effects (FE), μ_{ij} is time-invariant country-pair random effects (RE), and ε_{ijt} is the residual error term; more on this below.

The first econometric issue arises from the fact that the multilateral resistance (MR) factors (P_i and P_j) are unobservable. Hence, these Ps become imbedded in the regression's error term, while being correlated with the included variables in (5); see Anderson and van Wincoop (2003: equation 21). Given that the error term is no longer independent of the regressors, the estimates of (5) are biased. Anderson and van Wincoop correct for this bias by estimating "n" non-linear equations. The procedure is very cumbersome and applicable when "n" is very small. Baldwin and Taglioni (2007) have proposed using time-varying FE, x_{it} and x_{jt}, and time-invariant FE, μ_{ij}, for panel data, since the first two remove the time-series correlation and the third the cross-sectional correlation. But this alternative has three drawbacks. The first is that the number of country dummies is unmanageable (11,900 for 100 countries and 10 years). The second is that serial correlation is only partially removed by μ_{ij}; nor can one utilize time-varying μ_{ijt} because the number of country dummies would catch up with the sample size. The third is that even though μ_{ij} FE could yield unbiased estimates, the researcher would not be able to estimate the impact of important time-invariant variables (distance, RTA etc.) on X. For this reason, Carrère (2006, 231) proposes that μ_{ij} be treated as RE. Our conclusion is that the best practical solution to the MR problem is the specification given in (5); see Fratianni and Marchionne (2008) and Fratianni and Oh (2009).

The GE regressors --our second econometric issue-- can be correlated also with other unobservables, such as cultural and historical affinities, that may influence the propensity of country-pairs to trade. If omitted, these variables will be captured by the error term, which, in turn, will be correlated with income and RTAs (Baier & Bergstrand, 2007). Carrère (2006: 232) corrects for this bias by using the Hausman-Taylor instrumental variable technique.

The third issue deals with the serial correlation properties of the variables that enter the GE. Zwinkels and Beugelsdijk (2010) make the point that the GE applied on panel data may be subject to biased estimates because trade and FDI flows and explanatory variables like income are statistically non-stationary and potentially co-integrated; if co-integration cannot be rejected, then the GE must either be estimated in first differences or as a Vector-Error Correction Model (VECM). For the authors, the VECM method is to be preferred to the first-differenced GE because it is consistent with the theories underlying the GE.

The fourth and final issue tackles the GE specification form, namely whether a nonlinear GE is superior to a log-linear GE. Santos-Silva and Tenreyro (2006) argue that a specification like (5) presents two problems. The first is that the variance of error term depends on the values of the regressors, implying that its expected value will also depend on the regressors. The second is that the presence of '0' in trade and FDI data and the fact that the log of zero is undefined make the data sample not randomly drawn. The authors recommend that the GE be estimated with a Poisson pseudo-maximum likelihood.

BEST PRACTICE

In this section we estimate a parsimonious equation (5) for trade and FDI flows and focus on those best econometric practices discussed in the previous section. We use yearly bilateral trade for 177 countries and outward FDI flows from 30 OECD countries to 30 OECD countries and 102 non-OECD countries; both datasets cover the period 1985-2005. World Trade Analyzer by Statistics Canada is the source of trade data; the International Direct Investment Statistics by OECD is the source of FDI data. FDI data explicitly distinguish between zero and missing values. Since '0' may also mean that FDI is too small to record (the observation unit is $1,000,000), truncation is a potential statistical issue. Our FDI dataset consists of 15,518 non-zero observations and 9,770 zero-values. Trade data do not permit this distinction. A '0' was entered in the trade data when a country-pair had FDI, otherwise the observation was considered missing. Trade data truncation is less problematic than FDI data truncation because the observation unit is $1,000. Our trade dataset consists of 199,129 non-zero observations and 12,429 zero-values. Our independent variables include nominal incomes, geographic distance, cultural affinities captured by common language and shared land border (adjacency), and institutional affinities captured by RTAs.[4] Estimates for export flows are shown in Table 1 and those for outward FDI in Table 2.

[Insert Tables 1 and 2 Here]

Column one of the two tables tests (5) for year-FE, a common practice in panel studies. The regression explains 62% of the variance of trade and 47% of the variance of FDI. All coefficients have the expected signs, except for adjacency in the FDI regression, and are statistically significant at the 1% level. Column two adds country-FE; for example, see Clougherty and Grajek (2008) and Cuervo-

Cazurra (2008). This regression explains 14% more of the trade variance (42% of FDI) than regression one. Income and distance elasticities decline; the impact of RTA falls in the trade equation and becomes marginally significant in the FDI equation. Column three uses year-FE and country-pair RE to capture cross-sectional variation. The log-likelihood improves with respect to columns one and two. Column four combines FE of column two with RE of column three and represents our best practice to solve the MR problem. We call it multilateral-RE. The log-likelihood improves further; in the trade equation all coefficients are strongly significant, whereas in the FDI equation adjacency and RTA are not.

Column five checks for endogeneity of RTA formation and income (Carrère, 2006) by using the Hausman-Taylor instrumental variable technique. The inference from the Hausman-Taylor test suggests no substantial difference with respect to the multilateral-RE specification. The latter appears to be robust with respect to other potential endogeneity biases.

Columns six and seven deal with the implications of non-stationary time series. We follow Zwinkels and Beugelsdijk's VECM two-step procedure. In the first step, exports or FDI are regressed on the co-integrated incomes, and in the second step the residuals from the first step are regressed on the co-integrated equation, including MR factors. Again, we find no substantive difference between the multilateral-RE specification and the VECM specification.

The last column of the two tables deals with the zero-value problem. We recall that zeros are a distinct feature of both datasets, but in particular of the FDI dataset. Our specification corrects for multilateral resistance.[5] In the FDI Poisson estimation, all coefficients are statistically highly significant, with adjacency and RTA now showing quantitative effects not different from those in the trade equation. In the trade Poisson estimation, the coefficients are not substantively different from those of the trade multilateral-RE, except for a lower distance elasticity, a result that accords with Santos-Silva and Tenreyro's (2006). In sum, the Poisson estimation appears to correct for the distortion due to the presence of zero-value observations.

CONCLUSION

This note has discussed theoretical and statistical issues that are of direct concern to economists and international business scholars who apply GE in their research. We have suggested best practices that emanate from recent theoretical and empirical developments and can be summarized as follows: use country and year fixed effects with country-pair random effects regression estimation in the absence of significant zero-values in the dataset; use a Poisson pseudo-maximum likelihood estimation with country and year fixed effects and country-pair random effects in the presence of large number of zero-values in the dataset. Apart from the zero-value problem, the most important methodological message of this note is that GE should correct for multilateral factors.

REFERENCES

Anderson, J.E. 1979. A theoretical foundation for the gravity equation. *American Economic Review*, 69(1):106-115.

Anderson, J.E., & van Wincoop, E. 2003. Gravity with gravitas: A solution to the border problem. *American Economic Review*, 93(1):170-192.

Baier, S.L., & Bergstrand, J.H. 2007. Do free trade agreements actually increase members' international trade? *Journal of International Economics*, 71(1):72-95.

Baldwin, R., & Taglioni, D. 2007. Trade effects of the Euro: A comparison of estimators. *Journal of Economic Integration*, 22(4):780–818

Bergstrand, J.H. 1985. The gravity equation in international trade: Some microeconomic foundations and empirical evidence. *Review of Economics and Statistics*, 67(3):474-481.

Bergstrand, J.H. 1989. The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade. *Review of Economics and Statistics,* 71(1):143-153.

Brainard, L.S. 1997. An empirical assessment of the proximity-concentration trade-off between multinational sales and trade. *American Economic Review*, 87(4):520-544.

Buckley, P.J., Clegg, L.J., Cross, A.R., Liu, X., Voss, H., & Zheng, P. 2007. The determinants of Chinese outward foreign direct investment. *Journal of International Business Studies*, 38(4):499-518.

Carrère, C. 2006. Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model, *European Economics Review*, 50(2):223–247.

Clougherty, J., & Grajek, M. 2008. The impact of ISO 9000 diffusion on trade and FDI: A new institutional analysis. *Journal of International Business Studies*, 39(4):613–633.

Cuervo-Cazurra, A. 2008. The effectiveness of law against bribery abroad. *Journal of International Business Studies*, 39(4):634–651.

Deardorff, A. 1998. Determinants of bilateral trade: Does gravity work in a Neoclassical world? In J.A. Frankel (Ed.), *The Regionalization of the World Economy*:7-32, Chicago: University of Chicago Press.

Eden, L., & Rodriguez, P. 2004. How weak are the signals? International price indices and multinational enterprises. *Journal of International Business Studies*, 35(1):61-74.

Egger, P., & Pfaffermayr, M. 2004. Distance, trade and FDI: A Hausman-Taylor SUR approach. *Journal of Applied Econometrics*, 19(2):227-246.

Evenett, S.J. & Keller, W. 2002. On theories explaining the success of the gravity equation. *The Journal of Political Economy*, 110(2):281-316.

Fratianni, M.U. 2009. The gravity equation in international trade. In A. Rugman (Ed.), *The Oxford handbook of international business*: 72-89. New York: Oxford University Press.

Fratianni, M.U., & Marchionne, F. 2008. Heterogeneity in Trade Costs. *Economics Bulletin*, 6 (48): 1-14.

Fratianni, M.U., & Oh, C.H. 2009. Expanding RTA, trade flows, and the multinational enterprise. *Journal of International Business Studies*, 40(7):1206-1227.

Haveman, J., & Hummels, D. 2004. Alternative hypotheses and the volume of trade: The gravity equation and the extent of specialization. *Canadian Journal of Economics,* 37(1):199-218.

Hejazi, W., & Pauly, P.H. 2005. How do regional trade agreements affect intra-regional and inter-regional FDI? In L. Eden & W. Dobson (Eds.), *Governance, multinational and growth*:176-208. Cheltenham: Edward Elgar.

Helpman, E. 1984. A simple theory of trade with multinational corporations. *Journal of Political Economy*, 92(3):451-471.

Helpman, E. 1987. Imperfect competition and international trade: Evidence from fourteen industrial countries. *Journal of the Japanese and International Economies*, 1(1):62-81.

Helpman, E., & Krugman, P.R. 1985. *Market structure and foreign trade. Increasing returns, imperfect competition, and the international economy*. Cambridge, MA: MIT Press.

Kleinert, J., & Toubal, F. 2010. Gravity for FDI. *Review of International Economics*, 18(1):1-13.

Li, Q., & Vashchilko, T. 2009. Dyadic military conflict, security alliances, and bilateral FDI flows. *Journal of International Business Studies*, doi:10.1057/jibs.2009.91.

Markusen, J.R. 1984. Multinationals, multi-plant economies, and the gains from trade. *Journal of International Economics*, 16(3/4):205-226.

Markusen, J.R., & Venables, A.J. 1998. Multinational firms and the new trade theory. *Journal of International Economics*, 46(2):183-203.

Santos-Silva, J.M.C., & Terenyro, S. 2006. The log of gravity. *Review of Economics and Statistics*, 88(4):641-658.

Zwinkels, R.C.J. & Beugelsdijk, S. 2009. Gravity equations: Workhorse or Trojan horse in explaining trade and FDI patterns across time and space? *International Business Review*, 19(1):102-115.

Table 1 Trade GE: Exports

	One-Way	Three-Way	Pair-RE	Multilateral-RE	Hausman-Taylor	VECM		Poisson
	(1)	(2)	(3)	(4)	(5)	Step1 (6)	Step2 (7)	(8)
Importer's GDP	0.829***	0.486***	0.760***	0.557***	0.568***	0.796***		0.552***
Exporter's GDP	1.005***	0.482***	0.918***	0.498***	0.503***	0.953***		0.584***
Distance	-1.032***	-1.369***	-1.087***	-1.356***	-1.359***		-1.363***	-0.981***
Common Language	0.681***	0.691***	0.626***	0.679***	0.734***		0.681***	0.449***
Common Border	0.359***	0.271***	0.615***	0.525***	0.461***		0.524***	0.538***
RTA	1.220***	0.310***	0.400***	0.177***	0.116***		0.147***	0.254***
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Importer & Exporter-FE	No	Yes	No	Yes	Yes	No	Yes	Yes
Country-Pair Effects	No	No	RE	RE	RE	No	RE	RE
Constant	-27.08***	1.609	-23.25***	-1.034	-1.761**	-33.35***	12.23***	-10.69***
Over-Dispersion								1.206***
Observations	199,129	199,129	199,129	199,129	199,129	199,129	199,129	211,621
Pair Numbers			15,862	15,862	15,862		15,862	17,186
R-Square	0.617	0.703	0.614	0.699	0.702	0.525	0.169	0.003
Log-Likelihood	-424,131	-398,875	-368,217	-364,034		-445,619	-364,663	-2.1×10^{10}
LR-Test vs. One-Way		50511		120,193			118,935	
		(0.000)		(0.000)			(0.000)	
LR-Test vs. Three-Way				69,682				
				(0.000)				
BPLM-Test			369,993	206,118			204,840	
			(0.000)	(0.000)			(0.000)	

Note: p-values in parenthesis. LR is the Likelihood ratio. BPLM is ther Breusch-Pagan Lagrange multiplier. Robust standard errors clustered by pair method used. *p<0.1; **p<0.05; ***p<0.01.

Table 2 FDI GE: Outward FDI

	One-Way (1)	Three-Way (2)	Pair-RE (3)	Multilateral-RE (4)	Hausman-Taylor (5)	VECM Step1 (6)	VECM Step2 (7)	Poisson (8)
Importer's GDP	0.715***	0.328**	0.806***	0.749***	0.807***	0.647***		1.021***
Exporter's GDP	0.705***	0.437***	0.662***	0.572***	0.607***	0.734***		0.544***
Distance	-0.474***	-0.960***	-0.591***	-1.082***	-1.061***		-1.083***	-1.379***
Common Language	1.568***	0.826***	1.612***	0.753***	0.796***		0.755***	1.571***
Common Border	0.106	-0.0496	0.375*	0.117	0.0458		0.115	0.680***
RTA	0.880***	0.164*	0.265***	0.0843	0.0517		0.0808	0.281***
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Importer & Exporter-FE	No	Yes	No	Yes	Yes	No	Yes	Yes
Country-Pair Effects	No	No	RE	RE	RE	No	RE	RE
Constant	-30.47***	-5.740	-31.64***	-20.86***	-23.70***	-32.74***	10.12***	-25.49***
Over-Dispersion								0.915***
Observations	15,518	15,518	15,518	15,518	15,518	15,518	15,518	25,288
Pair Numbers			2,152	2,152	2,152		2,152	3,589
R-Square	0.469	0.670	0.452	0.659	0.666	0.364	0.305	0.041
Log-Likelihood	-32,718	-29,026	-28,024	-27,184		-34,125	-27,189	-4,735,000
LR-Test vs. One-Way		7382		11,067			11,058	
		(0.000)		(0.000)			(0.000)	
LR-Test vs. Three-Way				3,685			3,676	
				(0.000)			(0.000)	
BPLM-Test			17,877	5,485			5,536	
			(0.000)	(0.000)			(0.000)	

Note: See Table 1.

[1] On these developments, see Fratianni (2009).

[2] FOB price is the general free-alongside-ship export price; CIF is the pre-tariff or water's edge price of delivering a product to the importing country's entry (Eden & Rodriquez, 2004).

[3] For theories of horizontal and vertical MNC, see Helpman (1984), Markusen (1984), and Markusen and Venables (1998).

[4] See Fratianni and Oh (2009) for a definition of the RTAs.

[5] A word of caution: a Poisson estimation presents its own difficulties, such as incompatibility with many dummies, convergence, over-dispersion, and sensitivity to initial conditions.