Società Italiana degli Economisti
52.ma Riunione Scientifica Annuale
Università degli Studi di Roma Tre, Roma, 14-15 ottobre 2011

Trade costs in bilateral trade flows: Heterogeneity and zeroes in structural gravity models

Rosa Bernardini Papalia[1], Silvia Bertarelli[2]

THIS VERSION: SEPTEMBER 2011

ABSTRACT

This paper contributes to the explanation of international trade flows with structural gravity models taking heterogeneity and excess zeroes into account. We introduce a more general hypothesis on the structure of trade costs in Helpman et al. (2008) theoretical model that is capable of explaining over-dispersion in trade data. Zero inflated negative binomial models are considered to analyze the impact of trade costs, measured in terms of geographical distance and contiguity effects. An analysis related to a sample of 37 countries trade flows, with heterogeneous effects across sectors and trade integrated areas, such as APEC and EU, is presented. The size of exporting and destination economies, cultural and institutional factors are considered as influencing both the extensive and the intensive margin of trade.

Keywords: gravity models, heterogeneous firms, zero-inflated negative binomial estimator

JEL classification: F12, F14, F15.

[1] Dipartimento di Scienze Statistiche, Università di Bologna,Via Belle Arti, 41 - 40126 Bologna, tel.: +39 051 2098275, fax: +39 051 232153, e-mail: rossella.bernardini@unibo.it

[2] Dipartimento di Economia Istituzioni Territorio, Università di Ferrara, Via Voltapaletto, 11 – 44100 Ferrara, tel.: +39 0532 455065, fax: +39 0532 245951, e-mail: silvia.bertarelli@unife.it

1. Introduction

A large empirical literature using firm and plant-level data has documented the presence of firm's heterogeneity in productivity and fixed costs, implying productivity thresholds for firms IN ORDER to be capable to export[1]. An important corollary is that trade is determined along an extensive (number of firms) as well as an intensive (average exports per firm) margin. The extensive margin exists because less-productive firms that cannot cover their fixed costs will not export at all. Traditional theory generating gravity models (e.g. Anderson and van Wincoop, 2003) assume homogeneous firms and love for variety in consumption to ensure that all goods are traded everywhere, therefore they do not show any extensive margin and all changes occur along the intensive margin. To rationalize zeroes in trade flows, fixed costs of export are to be assumed as in Helpman et al. (2008), Hallak (2010) and Baldwin and Harrigan (2011). The selection mechanism due to productivity heterogeneity among firms determines which firms export in a country, how much goods are supplied by each exporter in foreign markets and therefore the amount of each country's aggregate exports.

This paper contributes to the explanation of international trade flows with structural gravity models taking heterogeneity and excess zeroes into account. We show how a gravity relationship with over-dispersion in trade flows arises from an extended version of Helpman et al. (2008) structural trade model in which unobserved heterogeneity in trade costs is explicitly modeled. This extension dictates a negative binomial rather than a Poisson estimator and zero-inflated models. Our estimation approach builds on Santos Silva and Tenreyro (2006) and Burger et al. (2009).

The paper is organized as follows. In section 2 a generalized version of Helpman et al. (2008) model with over-dispersion in trade costs is proposed and the corresponding econometric specification with excess zeroes and over-dispersion is presented in section 3. Section 4 reports evidence on a sample of 36 countries from NBER-UN database (Feenstra et al., 2005) and section 5 concludes with some final comments.

2. Theory

A gravity equation can be derived from models with firms' heterogeneity. For the purpose of our paper, we report the model developed by Helpman et al. (2008) that considers the self-selection process of firms on the basis of their productivity levels, and the influence that they have on international trade flows. This model is also capable to account for the probability that two countries may trade or not. This model therefore includes all possible combinations: i) zero values of the bilateral flows between country j and country i; ii) positive flows in one direction (e.g. from j

[1] See Greenaway and Kneller (2007) for a recent survey of the micro econometric evidence.

to i) and no flows in the other one (e.g. from i to j); iii) positive flows in both directions.
The authors consider a number of countries J, indexed by $j = 1, 2, ..., J$, and an indeterminate group of final goods, whose markets operate in a monopolistic competition framework. Each country includes N_j firms, each of which produces a single differentiated good. It follows that there are $N = \sum_{j=1}^{J} N_j$ products worldwide.
Any country j 's firm produces output at a cost $c_j a$, where a measures the amount of input required for the production of a unit of output, and c_j is the unit cost of inputs. The cost c_j varies across countries, reflecting international differences in production costs, and a across firms, thereby reflecting differences in productivity levels among firms even within the same country. The cumulative distribution of these levels of efficiency is represented by the function $G(a)$, identical in all countries, in which the two extremes are defined as $a_H > a_L > 0$.
The selling price of good l from country j is given by the cost of production multiplied by the mark-up, $1/q$:

[1] $$p_j(l) = \frac{c_j a}{q}$$

International trade implies that a firm located in country j sells its product in country i and incurs a fixed cost of access, $c_j f_{ij}$, and a transport cost equal to τ_{ij}. For the latter cost, the "melting iceberg" specification is assumed, that is τ_{ij} units of a product have to be shipped from country j to i for one unit to arrive. We assume that $f_{jj} = 0$ for every j and $f_{ij} > 0$ for $i \neq j$, and $\tau_{jj} = 1$ for every j and $\tau_{ij} > 1$ for $i \neq j$. Note that the fixed cost coefficients f_{ij} and the transport cost coefficients τ_{ij} depend on the identity of the importing and exporting countries, but not on the identity of the exporting producer. In particular, they do not depend on the producer's productivity level.
The selling price of good l from country j to country i is therefore equal to

[2] $$\check{p}_j(l) = \tau_{ij} \frac{c_j a}{q}$$

Using [2] we can define the quantity demanded by country i with the following equation:

[3] $$x_i(l) = \frac{\check{p}_j(l)^{1-\varepsilon} Y_i}{P_i^{1-\varepsilon}}$$

where Y_i identifies country i's income, and P_i its price index.
Using [2] and [3] we can determine the profit arising from j 's sales to i:

[4] $$\pi_{ij}(a) = (1-q)\left(\tau_{ij} \frac{c_j a}{q P_i}\right)^{1-\varepsilon} Y_i - c_i f_{ij}$$

The level of efficiency, $1/a_{ij}$, is derived from the zero profit condition and indicates the threshold above which a firm can achieve positive profits and then it is profitable to sell its product to foreign consumers:

[5] $$(1-q)\left(\tau_{ij} \frac{c_j a}{q P_i}\right)^{1-\varepsilon} Y_i = c_i f_{ij}$$

The zero profit condition [5] shows that only the fraction $G(a_{ij})$ of the N_j firms of country j export to country i. We now proceed to the determination of bilateral flows.

Given the price function [2] and the demand function [3], we can derive the value of imports of country i from country j:

[6] $$M_{ij} = \left(\frac{c_j \tau_{ij}}{qP_i}\right)^{1-\varepsilon} Y_i N_j V_{ij}$$

where

[7] $$V_{ij} = \begin{cases} \int_{a_L}^{a_{ij}} a^{1-\varepsilon} dG(a) & if\ a_{ij} \geq a_L \\ 0 & otherwise \end{cases}$$

As in Helpman et al. (2008) we assume that firm productivity $1/a$ is Pareto distributed, truncated to the support $[a_L, a_H]$. In this case we have

[8] $$G(a) = \frac{a^k - a_L^k}{a_H^k - a_L^k}, k > (\varepsilon - 1).$$

Then expression [7] becomes

[9] $$V_{ij} = \frac{k a_L^{k-\varepsilon+1}}{(k-\varepsilon+1)(a_H^k - a_L^k)} W_{ij}$$

where W_{ij} represents the firms' export share from country j to country i and is defined as follows:

[10] $$W_{ij} = max\left\{\left(\frac{a_{ij}}{a_L}\right)^{k-\varepsilon+1} - 1, 0\right\}$$

In this way, Helpman et al. (2008) obtain a gravity model, that includes the process of firm selection through the value of V_{ij}, to study the influence of costs and productivity levels on trade flows. The selection process of companies in foreign markets, represented by the variable W_{ij}, is then determined by the value of the "cutoff" a_{ij}, obtained from the zero profit condition in equation [5]. The trade equation can be written as

[11] $$M_{ij} = AA_i A_j \tau_{ij}^{1-\varepsilon} max\left\{\left[\frac{a_{ij}}{a_L}\right]^{k-\varepsilon+1} - 1, 0\right\}$$

Information on a_{ij} and a_L is typically not available. To overcome this problem, Helpman et al. (2008) define a latent variable Z_{ij}, which represents the value of the more productive firm's profit (with productivity $1/a_L$) compared to the fixed cost of exporting from j to i:

[12] $$Z_{ij} = \frac{(1-q)\left(\frac{qP_i}{c_j \tau_{ij}}\right)^{1-\varepsilon} Y_i a_L^{1-\varepsilon}}{c_i f_{ij}}$$

Trade flows are positive when $Z_{ij} > 1$. In this case, given equations [5] and [8], W_{ij} is a monotone function of Z_{ij}:

[13] $$W_{ij} = Z_{ij}^{(k-\varepsilon+1)/(\varepsilon-1)} - 1$$

By assuming i.i.d normal distributed errors on the random components of the model and by defining $T_{ij} = 1$ if country j exports to i (and $T_{ij} = 0$ otherwise) they obtain the probability to have export flows from j to i as a probit model and they estimate the trade equation for the positive observations of Mij, by using a procedure to account for the sample-selection issue. We instead propose to assume a more general form for random trade cost components, with the objective of modelling unobserved heterogeneity in trade flows across country pairs.

Specifically, we assume that the unobserved components of transport cost τ_{ij} and fixed cost to have access to foreign markets f_{ij} are gamma distributed. The transport cost depends on the geographical distance D_{ij} from i to j

[14] $$\tau_{ij} = D_{ij}^{\gamma} e^{-u_{ij}}$$

and on the u_{ij} disturbance, which is distributed i.i.d. gamma.

In this case trade equation [11] can be re-written as

$$M_{ij} = AA_i A_j D_{ij}^{-\gamma} max\left\{\left[\frac{a_{ij}}{a_L}\right]^{k-\varepsilon+1} - 1{,}0\right\} e^{u_{ij}}$$

To model the latent variable Z_{ij}, we assume that fixed costs f_{ij} are stochastic due to frictions or impediments related to trade. They depend on trade barriers imposed by the importing country to all countries (IM_j), on export-related fixed costs (EX_i) and on any additional fixed cost of each specific country pair (χ_{ij}). Specifically,

[15] $$f_{ij} = EX_j^{\zeta} IM_i^{\varsigma} \chi_{ij}^{k} e^{-v_{ij}}$$

Differently from other contributions, we assume that the unobserved residual v_{ij} is distributed i.i.d. gamma as transportation costs.

Given trade equation [11] and fixed costs equation [15], the latent variable Z_{ij} can be expressed as follows:

[16] $$Z_{ij} = BB_i B_j \chi_{ij}^{-k} D_{ij}^{-\gamma} e^{\eta_{ij}}$$

where

$$B \equiv (1-q) q^{1-\varepsilon} a_L^{1-\varepsilon} \ ,$$

$$B_i \equiv \frac{P_i^{1-\varepsilon} Y_i}{c_i IM_i^{\varsigma}} \ ,$$

$$B_j \equiv \frac{EX_j^{-\zeta}}{c_j}$$

and $\eta_{ij} \equiv u_{ij} + v_{ij}$, is gamma distributed when u_{ij} and v_{ij} are assumed to be independent[2]. In this case the density function of the error term is

[17] $$g(\eta_{ij}) = \frac{\delta^{\delta}}{\Gamma(\delta)} \eta_{ij}^{\delta-1} e^{-\eta_{ij}\delta}$$

[2] The gamma distribution has been used by Eaton et al. (2011) to model heterogeneity in a gravity model with a discrete choice structure. The authors exploit this characteristics with reference to the number of firms which trade internationally.

with $E(\eta_{ij})= (1)$ and $V(\eta_{ij})=1/\delta = \alpha$.

Given a consistent estimate of Z_{ij}, named $Z_{ij}^{*} = BB_i B_j \chi_{ij}^{-k} D_{ij}^{-\gamma}$, it is easy to show that the export flow M_{it} may be written as

[18] $$M_{ij} = T_{ij} AA_i A_j D_{ij}^{-\gamma} \left\{ \left[Z_{ij}^{*} e^{\eta_{ij}} \right]^{\frac{k-\varepsilon+1}{\varepsilon-1}} - 1 \right\} e^{u_{ij}}$$

with the indicator $T_{ij} = 1$ if country j exports to i and $T_{ij} = 0$ otherwise.

For the positive observations of M_{ij} we can take the logs of both sides of [18] to obtain

[19] $$m_{ij} = a + a_i + a_j - \gamma d_{ij} + \ln\left[e^{\delta\left(z_{ij}^{*}+\eta_{ij}\right)} - 1 \right] + u_{ij}$$

where $\delta = \frac{k-\varepsilon+1}{\varepsilon-1}$ and lower-case letters indicate the logs of the quantity corresponding to the same upper case letters.

3. The structural gravity models with over-dispersion and excess zeroes

The empirical literature on zeroes in bilateral trade data includes Eaton and Tamura (1994), Santos Silva and Tenreyro (2006), Helpman et al. (2008), and Eaton et al. (2011) among others. As presented in the previous section, the analysis of trade flows is to be corrected for the probability of positive trade in gravity models. Helpman et al. (2008) consider the Heckman selection model, where the selection equation determines whether or not bilateral trade between two countries is observed, while the regression model concentrates on the analysis of bilateral trade determinants. This model requires some restrictive assumptions of homoskedastic random components. Despite the fact that the Heckman selection model deals with zero counts, the bias created by the logarithmic transformation in the regression part of the model poses a problem in the presence of heteroskedasticity, even when controlling for fixed effects. Santos Silva and Tenreyro (2006) propose a Poisson pseudo-maximum likelihood (PPML) estimator to be applied to the gravity model in the multiplicative specification given by Helpman et al. (2008). However, they observe that this estimator does not take full account of the heteroskedasticity in the model and all inference has to be based on an Eicker-White robust covariance matrix operator[3].

In this view, we directly estimate the gravity model in the multiplicative specification, as in Santos Silva and Tenreyro (2009), to avoid the problems associated to Helpman et al. (2008) estimation procedure in the presence of heteroskedasticity, We also take into account the presence of zeroes in the data.

[3] Other contributions in estimating gravity models consider the PPML approach. See for example Westerlund and Wilhelmsson (2011) and Liu (2009).

With reference to trade equation [18] we can summarize all observables by the vector $X_{ij} = (a_i, a_j, b_i, b_j, ln\chi_{ij}, d_{ij})$, such as geographical distance, language, cultural and institutional variables. In addition, importer and exporter fixed effects respectively control for inward and outward multilateral resistance variables (Anderson and van Wincoop, 2003).

In this case, the conditional mean of the trade flow between country *i* and country *j* when $T_{ij} = 1$ can be written as a function of the vector of covariates and of an unobserved term for each (i, j) country pair:

[20] $$E(M_{ij}|X_{ij}, \eta_{ij}) = \mu_{ij}\eta_{ij} = exp^{(X'_{ij}\beta)}\eta_{ij}$$

Interestingly, given the gamma distribution for trade (fixed and variable) costs, we have the nice feature that if we do not condition on η_{ij} - which we cannot do in practice since it is not observed - the trade flow M_{ij} is distributed negative binomial. More specifically, conditional on both X_{ij} and η_{ij}, the dependent variable M_{ij} is Poisson distributed with probability

[21] $$\mathrm{P}(M_{ij} = I_{ij}|X_{ij}, \eta_{ij}) = \frac{e^{-\mu_{ij}\eta_{ij}}(\mu_{ij}\eta_{ij})^{I_{ij}}}{I_{ij}!}$$

However, conditional on only X_{ij}, M_{ij} is distributed as a negative binomial

[22] $$\mathrm{P}(M_{ij} = I_{ij}|X_{ij}) = \frac{\Gamma(I_{ij}+\alpha^{-1})}{I_{ij}!\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\mu_{ij}+\alpha^{-1}}\right)^{\alpha^{-1}}\left(\frac{\mu_{ij}}{\mu_{ij}+\alpha^{-1}}\right)^{I_{ij}}$$

where $E(M_{ij}) = \mu_{ij}$ and $V(M_{ij}) = \mu_{ij}(1+\alpha\,\mu_{ij})$.

Negative binomial and Poisson models are nested because the negative binomial converges to Poisson, as α converges to 0. The negative binomial distribution tests for the presence of *unobserved* heterogeneity in the sample, while the Poisson model assumes equi-dispersion and therefore is able to account for *observed* heterogeneity only. If the estimation procedure does not correct for over/under-dispersion, results are consistent but inefficient (with spuriously large z-values and small p-values due to downward standard errors).

The negative binomial distribution emerges when M_{ij} is strictly positive. However, no bilateral trade can exist as well, then both $T_{ij} = 1$ and $T_{ij} = 0$ alternatives must be considered in our estimation strategy. This can be done by considering a zero-inflated version of the negative binomial model.

The zero-inflated model considers two kinds of zero-valued trade flows: countries that never trade and countries that do not trade now but potentially could trade in the future, based on a positive latent probability to trade obtained by some determinants such as distance, institutional proximity, etc. In this view, this model accounts for unobserved heterogeneity in the population with a zero count.

Then the estimation process consists of two parts. The first equation contains a logit (or probit) regression of the probability that there is no bilateral trade at all. With reference to the structural model presented in section 2, no trade occurs when $Z_{ij} \leq 1$. The second part contains a negative binomial regression of the probability of each count for the group that has a non zero probability or interaction intensity other than zero.

Formally we have:

[23] $$Pr[T_{ij}=0] = (1-\chi_{it}) + \chi_{it}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{ij}}\right)^{\alpha^{-1}}$$

[24] $$Pr[M_{ij}=I_{ij} \mid T_{ij}=1] = \chi_{it}\ \frac{\Gamma(I_{ij}+\alpha^{-1})}{I_{ij}!\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{ij}}\right)^{\alpha^{-1}}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{ij}}\right)^{I_{ij}}$$

where the mean value μ_{ij} has been defined in equation [20].

Finally, further heterogeneity may be captured by admitting differences in trade cost coefficients across industries (Baldwin and Taglioni, 2006; Martínez-Zarzosoa and Pérez-Garcíab, 2008; Chen and Novy, 2008).

4. Estimating structural gravity models from sector data

The analysis of trade flows of 37 countries in 2000 has been performed by using the NBER-UN data set, described in Feenstra et al. (2005), for sectors in the 4-digit Standard International Trade Classification (SITC4), Revision 2. See table 2 for a list of countries. For each sector we consider 37×36 country pairs (i, j) trade flows where the number of zero observations is high (72% on total observations). To model trade costs, two variables connected to distance are considered. The geographical distance between the capital cities of each country pair captures variable trade costs, which are increasing with the distance. Its effect on trade volume is expected to be negative. A contiguity variable is connected to fixed trade costs and is equal to 1 when two countries share a common border. Its effect on trade volume is expected to be positive. Since it is argued that the geographical distance is not able to capture all the economics barriers to trade, the analysis often consider other gravity variables. As to other explanatory variables, we have: a language dummy, which is equal to one when two countries share the same language; a history dummy which is equal to one when two countries had a colonial link or were the same country in the past. All these variables are from CEPII. Country i's log GDP and country j's log GDP (Source: Penn Table) are used to consider trading partners' dimensions; four free trade agreement dummies are equal to one when two countries join the same free trade area (APEC, EFTA, EU, NAFTA). To model multilateral resistance effects, exporter fixed effects and importer fixed effects have been introduced.

Given that omitting the zero values gives biased results, a zero-inflated negative binomial estimator has been considered to account for both the extensive margin, expressed in terms of the probability to have a zero trade flow between country pairs, and the intensive margin of trade.

Results, presented in table 2, are compared by considering two alternatives. The first one considers cluster-robust OLS estimates with reference to the logarithmic variable ln(Value) for which zero observations are dropped, and the second one refers to the logarithmic variable ln(Value+0.01) for which zero observations are maintained. With reference to OLS estimates, most variables have the expected sign and are highly statistically significant for the log-normal specification. Trade decreases with geographical distance, and increases when countries share the same language and with their size. However, the coefficients of the determinants of trade in the OLS gravity equation are biased, as they confound the effect of these variables on the intensive margin of trade with their indirect effect on the probability to have positive export flows. The results of the estimation of the log-normal model with zero valued flows show the same signs, but differ substantially in size (50% or more).

The zero-inflated model generates two sets of parameter estimates (both reported in table 2): one set for the logit model, which identifies members of the group of pairs of countries always having zero values (pairs of countries that never trade), and one set for the negative binomial part, which predict the probability of a count belonging to the group of countries that have theoretically non-zero trade flows. The correct model choice depends on the extent to which over-dispersion and excess zeroes are empirically relevant. To test the null hypothesis of the α parameter of over-dispersion equal to zero a likelihood ratio test has been performed rejecting the null in all specifications. With reference to excess zeroes, a Vuong test shows that a positive value in favor of the zero-inflated model cannot be rejected.

As to the intensive margin of trade, the geographical distance coefficient is negative, while the size of the trading economies, the common language, the same free trade area, the contiguity, and the history variable have positive effects. The sign of coefficients in the logit model, which considers the probability to have no trade flows, present as expected an opposite sign with respect to intensive margin equation ones. Table 3 reports exporter and importer fixed effects as a measure of outward and inward multilateral resistance terms, with USA as the benchmark country (orange coefficients are not significantly different from US term), along an increasing order. When estimating the same model with heterogeneous coefficients across APEC and EU countries (table 4), we observe that while geographical distance has similar negative effects on trade intensity, contiguity affects trade intensity within APEC area in a stronger way than within EU and past colonial ties are not relevant for APEC countries.

By admitting heterogeneity across sectors in slope coefficients, we observe that the assumption of homogeneous effects of trade costs, captured by distance and contiguity variables, cannot be accepted (see tables 5-9). Moreover, interesting differences in the multilateral resistance terms emerge. Our findings are in line with previous evidence on industry level data, such as Baldwin and Taglioni (2006), Chen and Novy (2008), Martínez-Zarzosoa and Pérez-Garcíab (2008).

5. Concluding remarks

This paper has contributed to the explanation of international trade flows by the specification of a structural gravity model taking heterogeneity and excess zeroes into account. We have introduced a more general hypothesis on the structure of trade costs in the Helpman et al. (2008) theoretical model that is capable of explaining over-dispersion in trade data in addition to zero trade flows between some pairs of countries and larger numbers of exporters to larger destination markets. Empirically, we have applied zero inflated negative binomial models to a sample of 37 countries in 2000 with the objective to explain the impact of trade costs measured in terms of geographical distance and contiguity effects. The size of exporting and destination economies, cultural and institutional factors such as common language, free trade agreement and colonial ties have been considered in influencing both the extensive and the intensive margin of trade. Heterogeneous trade costs characteristics emerge across sectors and across trade integrated areas, such as APEC and EU.

References

Anderson, James, and Eric van Wincoop (2003) Gravity with gravitas: a solution to the border puzzle, *American Economic Review* 93, 170–92.

Baldwin R.E. Harrigan J. (2011) Zeros, quality and space: trade theory and trade evidence, *American Economic Journal: Microeconomics* 3 (May 2011), 60-88.

Baldwin R.E. Taglioni D. (2006) Gravity for Dummies and Dummies for Gravity Equations, NBER Working Paper 12516.

Burger M.J. Van Oort F. Linders G.M. (2009) On the specification of the gravity model of trade: zeros, excess zeros and zero-inflated estimation, *Spatial Economic Analysis*, 4, 167–190.

Chen N. and Novy D. (2008) Gravity, trade integration and heterogeneity across industries, CEPR working paper n. 7103.

Eaton J. Kortum S. Sotelo S. (2011) International Trade: Linking micro and macro, mimeo. Available at http://home.uchicago.edu/~kortum/papers/EKS_20_02_2011_tabfig.pdf .

Feenstra Robert C. Lipsey R.E. Deng H. Ma A. Mo H. (2005) *World trade flows: 1962-2000*, NBER working paper n. 11040.

Greenaway, David and Richard Kneller (2007) Firm Heterogeneity, Exporting and Foreign Direct Investment, *Economic Journal*, Vol 117, F134-F161.

Hallak J.C. (2010) A product-quality view of the Linder hypothesis, *Review of Economics and Statistics* 92(3), 453-466.
Helpman E. Melitz M. Rubinstein Y. (2008) Estimating trade flows: trading partners and trading volumes, *The Quarterly Journal of Economics*, 123 (2), 441-487.
Liu X. (2009) GATT/WTO promotes trade strongly: sample selection and model specification, *Review of International Economics* 17(3), 428-446.
Martínez-Zarzosoa I. Pérez-Garcíab E.M. (2008) Do transport costs have a differential effect on trade at the sectoral level? *Applied Economics* 40, 3145-57.
Santos Silva J.M.C. Tenreyro S. (2006) The log of gravity, *The Review of Economics and Statistics*, 88(4), 641-658.
Westerlund J. Wilhelmsson F. (2011) Estimating the gravity model without gravity using panel data, *Applied Economics* 43(6), 641-649.

Table 1: List of countries

EU members	APEC members
Austria	Australia
Belgium-Lux	Canada
Czech Rep	China
Denmark	Chile
Finland	Japan
France	Korea Rep.
Germany	Mexico
Greece	New Zealand
Hungary	Russian Federation
Ireland	Singapore
Italy	USA
Netherlands	
Poland	
Portugal	**Other countries**
Romania	Iceland
Slovakia	Israel
Slovenia	Norway
Spain	South Africa
Sweden	Switzerland
UK	Turkey

Table 2: Gravity estimates, homogeneous slope coefficients across sectors

	Robust OLS Ln(Value)	**Robust OLS** Ln(Value+0.01)	**Negative binomial** Extensive margin	**Negative binomial** Intensive margin
Geographical distance	-0.33***	-0.90***	0.63***	-0.28***
Contiguity dummy	0.50***	1.32***	-0.52***	0.52***
Language dummy	0.11***	0.52***	-0.34***	0.03**
History dummy	0.12***	0.44***	-0.33***	0.25***
GDP importer	0.28***	0.68***	-0.50***	0.52***
GDP exporter	0.29***	0.89***	-0.74***	0.23***
APEC dummy	0.41***	0.82***	-0.53***	0.75***
EFTA dummy	0.13**	0.11**	-0.37***	-0.30***
EU dummy	0.17***	0.20***	-0.11***	0.29***
NAFTA dummy	0.82***	1.96***	-0.57***	0.89***
Constant	-7.80***	-36.96***	33.40***	-7.97***

Estimates with importer, exporter and (1 digit SITC) sector fixed effects; *** 1%, ** 5%, * 10% significant coefficient. Parameter α = 2.56 (p-value: 0.00), Vuong test 209.12 (p-value: 0.00).

Table 3: Outward and inward multilateral resistance terms

Country	Outward	Country	Inward
Greece	-1.43	Russian Fed	-1.28
Slovenia	-1.42	China	-1.01
Poland	-1.34	Chile	-0.99
Czech Rep	-1.25	Romania	-0.95
Slovakia	-1.15	Slovakia	-0.79
Romania	-1.15	Poland	-0.77
Austria	-1.13	Czech Rep	-0.75
Turkey	-1.08	New Zealand	-0.73
Mexico	-1.03	Mexico	-0.71
Portugal	-1.00	Slovenia	-0.70
Hungary	-0.91	South Africa	-0.65
Canada	-0.89	Austria	-0.61
Spain	-0.84	Greece	-0.59
Denmark	-0.82	Canada	-0.59
Israel	-0.62	Hungary	-0.48
New Zealand	-0.61	Turkey	-0.47
Belgium	-0.55	Finland	-0.46
Netherlands	-0.53	Korea Rep.	-0.44
France	-0.52	Australia	-0.44
Switzerland	-0.49	Denmark	-0.34
China	-0.45	Portugal	-0.34
Italy	-0.44	Norway	-0.34
UK	-0.38	Spain	-0.33
Finland	-0.37	France	-0.29
South Africa	-0.33	Japan	-0.25
Australia	-0.33	Italy	-0.24
Sweden	-0.30	Sweden	-0.18
Korea Rep.	-0.23	Switzerland	-0.15
USA	0	Netherlands	-0.08
Chile	0.03	UK	-0.07
Norway	0.05	Israel	-0.04
Ireland	0.08	USA	0
Germany	0.08	Germany	0.07
Singapore	0.20	Ireland	0.08
Japan	0.24	Belgium	0.16
Russian Fed	0.59	Singapore	0.40

Table 4: Gravity estimates, heterogeneous slope coefficients across Free Trade Areas

	EU		APEC	
	Negative binomial Extensive margin	**Negative binomial** Intensive margin	**Negative binomial** Extensive margin	**Negative binomial** Intensive margin
Geographical distance	0.72***	-0.47***	0.94***	-0.42***
Contiguity dummy	-0.54***	0.39***	-0.54***	0.96***
Language dummy	-0.08**	0.11***	-0.55***	-0.24***
History dummy	-0.29***	0.19***	1.14***	-0.09
GDP importer	-0.66***	0.68***	-0.37***	0.68***
GDP exporter	-0.90***	0.63***	-0.69***	0.39***
Constant	41.46***	-23.41***	25.88***	-15.53***

Estimates with importer, exporter and (1 digit SITC) sector fixed effects; *** 1%, ** 5%, * 10% significant coefficient. EU: parameter $\alpha = 2.28$ (p-value: 0.00), Vuong test 137.15 (p-value: 0.00); APEC: parameter $\alpha = 3.11$ (p-value: 0.00), Vuong test 69.24 (p-value: 0.00).

Table 5: High tech sector gravity estimates

	Negative binomial Extensive margin	**Negative binomial** Intensive margin
Geographical distance	0.36***	-0.27***
Contiguity dummy	-0.58***	0.26***
Language dummy	-0.41***	0.17***
History dummy	-0.28***	0.24***
GDP importer	-0.49***	0.58***
GDP exporter	-0.84***	0.72***
APEC dummy	-0.65***	0.82***
EFTA dummy	-0.30*	-0.25
EU dummy	0.06	0.19***
NAFTA dummy	-0.80***	0.94***
Constant	34.44***	-25.94***

Estimates with importer, exporter and (1 digit SITC) sector fixed effects; *** 1%, ** 5%, * 10% significant coefficient. Parameter $\alpha = 2.42$ (p-value: 0.00), Vuong test 69.51 (p-value: 0.00).

Table 6: Outward and inward multilateral resistance terms, high tech sector

Country	Outward
Chile	-2.29
Russian Fed	-2.13
South Africa	-1.46
Poland	-1.28
Turkey	-0.92
Australia	-0.88
Mexico	-0.81
Spain	-0.79
China	-0.70
New Zealand	-0.67
Portugal	-0.66
Czech Rep	-0.62
Canada	-0.57
Italy	-0.49
Austria	-0.48
Greece	-0.43
Romania	-0.30
Slovakia	-0.27
Norway	-0.26
France	-0.11
Korea Rep.	0.00
USA	0
Japan	0.15
Belgium	0.23
UK	0.25
Germany	0.29
Denmark	0.43
Israel	0.59
Slovenia	0.62
Netherlands	0.65
Hungary	0.71
Sweden	1.10
Switzerland	1.12
Finland	1.25
Singapore	1.45
Ireland	1.62

Country	Inward
Russian Fed	-1.36
Chile	-1.36
China	-1.12
Romania	-1.09
Slovenia	-1.05
Slovakia	-0.92
Mexico	-0.89
Czech Rep	-0.78
Poland	-0.77
South Africa	-0.70
Canada	-0.69
Greece	-0.66
New Zealand	-0.64
Turkey	-0.62
Austria	-0.54
Finland	-0.51
Korea Rep.	-0.49
Norway	-0.48
Spain	-0.44
Israel	-0.39
Italy	-0.39
Australia	-0.37
Portugal	-0.37
Japan	-0.36
Denmark	-0.35
France	-0.27
Hungary	-0.13
Belgium	-0.10
Sweden	-0.07
Switzerland	-0.03
USA	0
Germany	0.03
UK	0.08
Netherlands	0.34
Singapore	0.56
Ireland	0.59

Table 7: Medium-high tech sector gravity estimates

	Negative binomial Extensive margin	**Negative binomial** Intensive margin
Geographical distance	0.67***	-0.40***
Contiguity dummy	-0.53***	0.42***
Language dummy	-0.38***	0.10***
History dummy	-0.32***	0.14***
GDP importer	-0.57***	0.54***
GDP exporter	-0.89***	0.53***
APEC dummy	-0.42***	0.55***
EFTA dummy	-0.41***	-0.15
EU dummy	-0.01	0.25***
NAFTA dummy	-0.64***	1.19***
Constant	37.48***	-18.74***

Some sub-sectors of chemicals, machinery and transport equipment
Estimates with importer, exporter and (1 digit SITC) sector fixed effects; *** 1%, ** 5%, * 10% significant coefficient.
Parameter α (p-value: 0.00), Vuong test (p-value: 0.00).

Table 8: Medium-low tech sector gravity estimates (without S3 sector)

	Negative binomial Extensive margin	**Negative binomial** Intensive margin
Geographical distance	0.64***	-0.30***
Contiguity dummy	-0.43***	0.62***
Language dummy	-0.28***	-0.13***
History dummy	-0.41***	0.31***
GDP importer	-0.50***	0.49***
GDP exporter	-0.71***	0.17***
APEC dummy	-0.67***	0.63***
EFTA dummy	-0.44***	-0.57***
EU dummy	-0.01	0.41***
NAFTA dummy	-0.44***	0.59***
Constant	32.01***	-6.34***

Estimates with importer, exporter and (1 digit SITC) sector fixed effects; *** 1%, ** 5%, * 10% significant coefficient.
Parameter α = (p-value: 0.00), Vuong test (p-value: 0.00).

Table 9: Low tech sector gravity estimates

	Negative binomial Extensive margin	**Negative binomial** Intensive margin
Geographical distance	0.66***	-0.28***
Contiguity dummy	-0.59***	0.60***
Language dummy	-0.39***	0.14***
History dummy	-0.31***	0.26***
GDP importer	-0.46***	0.48***
GDP exporter	-0.64***	0.10***
APEC dummy	-0.64***	0.88***
EFTA dummy	-0.42***	0.03
EU dummy	-0.27***	0.34***
NAFTA dummy	-0.63***	1.03***
Constant	28.62***	-5.20***

Estimates with importer, exporter and (1 digit SITC) sector fixed effects; *** 1%, ** 5%, * 10% significant coefficient. Parameter $\alpha =$ (p-value: 0.00), Vuong test (p-value: 0.00).

Table 10: Mineral fuels, lubricants and related material sector gravity estimates (S3)

	Negative binomial Extensive margin	**Negative binomial** Intensive margin
Geographical distance	1.01***	-0.86***
Contiguity dummy	-0.73***	0.39***
Language dummy	-0.12	0.06
History dummy	-0.36***	0.43***
GDP importer	-0.49***	0.42***
GDP exporter	-0.86***	0.53***
APEC dummy	-0.59***	0.95***
EFTA dummy	0.62	-1.09
EU dummy	-0.04	-0.42**
NAFTA dummy	0.19	-0.57*
Constant	34.22***	-12.71**