The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Capturing zero trade values in gravity equations of trade: A disaggregated sectoral analysis

George Philippidis[a], Helena Resano-Ezcaray[b], and Ana I. Sanjuán-López[b]
[a]Aragonese Agency for Research and Development (ARAID), Government of Aragón, Avda Montañana 930, 50059 Zaragoza, Spain
[b]Deparment of Agrofood Economics and Natural Resources, Centre of Agrofood Research and Technology of Aragon (CITA), Government of Aragón, Avda Montañana 930, 50059 Zaragoza, Spain

Paper prepared for presentation at the EAAE 2011 Congress
Change and Uncertainty
Challenges for Agriculture,
Food and Natural Resources

August 30 to September 2, 2011
ETH Zurich, Zurich, Switzerland

Copyright 2011 by [George Philippidis, Helena Resano-Ezcaray, and Ana I. Sanjuán-López]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

1. Introduction

The gravity method provides a benchmark for trade under frictionless conditions. In its simplest form, trade between a pair of countries is a positive function of economic 'size' (GDP) and a negative function of trade costs (proxied by distance). Since the early work of Tinbergen (1962) various authors (Anderson, 1979; Bergstrand, 1990; Deardorff, 1998) have theoretically grounded the approach of employing a homothetic constant elasticity substitution (CES) Armington (1969) structure consistent with the assumption of monopolistic competition (See chapter five of Feenstra (2004) for a comprehensive review of the theoretical and empirical development of the gravity equation).

From a methodological perspective, cross-section data cannot fully control for country heterogeneity, which typically biases the results and although country 'fixed effects' may solve this problem, it brings collinearity with theoretically relevant explanatory variables. For this reason, the use of panel data with 'fixed effects' dummies are employed to improve the estimation of the parameters of interest in the gravity equation (Westerlund and Wilhelmsson, 2006). The majority of recent gravity works (e.g. Burger et al., 2009; Wang et al., 2010) apply panel data to predict total trade between bilateral partners. In this study, although we employ a panel of only two years (2001 and 2004), an important development of this approach is that we estimate predicted gross bilateral trade flows in much greater sectoral detail (57 separate sectors) between 95 partner countries/regions. This study is restricted to the analysis of the 20 agro-food sectors.

A further development of this research centers on the application of different estimation methods. In general, gravity studies apply the ordinary least squares (OLS) specification, whilst more recent literature demonstrates that the Poisson Maximum Likelihood Estimator solves some of the typical drawbacks of the OLS model (heteroskedasticity and omission of zero trade flows). Notwithstanding, Poisson model and its variants (Negative Binomial, Zero Inflated Poisson and Zero Inflated negative Binomial) may still suffer from model misspecification issues leading to inconsistent estimates. In particular, the endogeneity of trade barriers (i.e., the 'endogeneity' problem) is still present, while omitted variables, such as the quantifiable existence of NTBs, may be leading to some bias in our parameter estimates.

To the best of our knowledge, there is only one other gravity application (Burger *et al.*, 2009) which handles the problems of heteroskedasticity and the omission of zero trade flows through the estimation of Poisson model and its variants, whilst with the exception of our study, there are no gravity studies which apply these methods on a sector-by-sector basis. Examining the results of our study, we discover that unlike previous attempts to dissect the partial impacts of tariffs and subsidies on bilateral trade for disaggregated sectors (Kuiper and van Tongeren, 2006; Philippidis et al., 2007), the estimated signs for both tariffs and export subsidies across the 20 sectors (with a very few exceptions) are consistent with *a priori* expectations and are statistically significant.

2. The theoretical model: Count models[1]

2.1 The Poisson and Negative Binomial models

The Poisson model belongs to the category of count data models in which the dependent variable is discrete, with a distribution that places probability mass at nonnegative integer

[1] An essential reference for count models is Cameron A.C. and Trivedi P.K. (1998).

values only. In the Poisson model, the occurrence of an event follows a Poisson distribution with density:

$$\Pr[Y = y_i] = P[\mu_i] = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \tag{1}$$

where μ is the rate parameter; the mean of y_i is $E[y_i] = \mu_i$ and the variance $V[y_i] = \mu_i$. The model is estimated by Maximum Likelihood.

The Poisson regression often suffers from the presence of over-dispersion. In other words, when the variance of the data exceeds the mean, this leads to deflated standard errors and grossly inflated t-statistics. Consequently, inference has to be based on White's (Eicker-White) robust covariance matrix estimator. Alternatively, a Negative Binomial Model can be used instead which relaxes the assumption of equi-dispersion of the Poisson model. The negative binomial (NB) improves the Poisson performance in the case of over-dispersion, and assumes as the conditional variance function: $Var[y_i] = \mu_i + \alpha\,\mu_i^2$, with α being the dispersion parameter and μ_i the conditional mean.

2.2. Zero Inflated Poisson and Zero Inflated Negative Binomial

In those cases where the dataset contains an excess of zeros, the zero inflated counterparts of Poisson and Binomial models, Zero Inflated Poisson (ZIP) and Zero Inflated Binomial (ZIB), are recommended. Both, ZIP and ZIB have been applied to a variety of fields including bio-diversity, medicine, manufacturing defects (Lambert, 1992) and economics (Cameron and Englin, 1997). However, their application within the agricultural economics literature is still rare, and to the best of our knowledge, has only been applied by Edmeades and Smale (2006) and Durham *et al.* (2004). Moreover, the application of zero inflated models within the gravity literature has been restricted to only one study (Burger *et al.,* 2009).

The panel data employed in this study reveals the presence of an excess of zeros in some sectors, and accordingly, the use of inflated models could be a logical methodological approach. In this context, the ZIB model handles over-dispersion by following a mixture of two states: one state leading to an excess of zeros, with probability p_i (i.e. probability of non-trade); and the second state, with probability $1\text{-}p_i$ (i.e. probability of trade). According to this approach, zeros may arise from: a 'zero-process' where only the 'true' zero values are observed; or a 'negative binomial process' that accounts for non-zero values and a proportion of 'false' zeros (Martin *et al.*, 2005). The ZIB model defines the following count probabilities (Shankar *et al.*, 1997; Greene, 2002):

$$\Pr(y_i = 0) = p_i + (1 - p_i)\left[\frac{\theta}{\theta + \lambda_i}\right]^{\theta} \tag{2}$$

$$\Pr(y_i = y) = (1 - p_i)\frac{\Gamma(\theta + y_i)u_i^{\theta}(1 - u_i)^k}{y_i!\,\Gamma(\theta)} \quad y>0 \tag{3}$$

Where θ is the inverse of the over-dispersion parameter α, θ = 1/α; $u_i = \theta/(\theta+\lambda_i)$ with λ_i the mean, and the expression $\frac{\Gamma(\theta + q_i)u_i^{\theta}(1\text{-}u_i)^k}{y_i!\Gamma(\theta)}$ is the negative binomial probability distribution.

When α equals to 1, the ZIB model converts into the ZIP model.

We use only one set of covariates, and accordingly, we are assuming that the same variables explain both, trade and non-trade (given the nature of our dataset, we cannot isolate the factors that affect each of these outcomes). Furthermore, this assumption simplifies the interpretation of the coefficients.

3. Empirical Application and Results
3.1 Final model specification

On the basis of the discussion in the preceding sections and following the work of Silva and Tenreyro (2006) and Burger *et al.* (2009), it seems clear that the Poisson model, as well as the Negative Binomial, and zero-inflated models are a more correct way to estimate the gravity equation because they obtain consistent estimates for the parameters of interest and account for zero trade values. For the purposes of estimating, we employ the software package STATA, which allows for identifying and removing those variables (i.e., fixed effects) which cause collinearity, which constitutes a considerable time saving for the modeller.

In our final specification we have tried to reconcile the theoretical basis of the gravity model with the feasibility of the estimation. The former implies the application of those explanatory variables consistent with theory (e.g., size of countries) and also those found relevant in empirical applications (e.g., remoteness), while the latter implies that multi-collinear variables need to be excluded. Nevertheless, as explained above, the problem of collinearity is mitigated thanks to the use of a panel of two years, and only in a few sectors was it found necessary to remove fixed effects.

The final estimated model is:

$$\begin{aligned} X_{ijt} = {} & \alpha + \beta_{mt}\, Mt_{ijt} + \beta_{xs}\, Xs_{ijt} + \beta_{Dist}\, Dist_{ij} + \beta_{Contig}\, Contig_{ij} + \\ & \beta_{Lang}\, Lang_{ij} + \beta_{Colony}\, Colony_{ij} + \beta_{NoSo}\, NoSo_{ij} + \beta_{Remote}\, Remote_{ijt} + \\ & \beta_{FTA}\, FTA_{ij} + \beta_{Gdp}\, Gdp_{ijt} + \beta_{Sqincome}\, Sqincome_{ijt} + \beta_{Year2004}\, Year2004_{t} + \\ & \sum_{i=2}^{T} \theta_i F_i + \sum_{j=2}^{T} \theta_j F_j + \varepsilon_{ijt} \end{aligned} \tag{4}$$

where:

X_{ijt} is the value of imports from country *j* on exports from *i* at world prices in year *t*,

Mt_{ijt} is the power of the import tariff rate applied by importer j on exports from *i* in year *t*, measured in *ad-valorem* equivalents, in logs: $\ln\left(1+\frac{AdvRate_{ijt}}{100}\right)$;

Xs_{ijt} is the *ad-valorem* power of the export subsidy rate applied by country *i* on exports to country *j* in year *t*, in logs: $\ln\left(1+\frac{XSubRate_{ijt}}{100}\right)$;

$Dist_{ij}$ is the great circle distance between the capitals of country *i* and *j* in logs;

$Contig_{ij}$ is a dummy variable that values 1 when countries *i* and *j* share a border, and 0 otherwise;

$Lang_{ij}$ is a dummy variable that values 1 when countries *i* and *j* share the same official language, and 0 otherwise;

$Colony_{ij}$ is a dummy variable that values 1 when countries *i* and *j* have or have had a colonial linkage;

$NoSo_{ij}$ is the difference in latitudes between countries *i* and j: $\ln(latitude_i - latitude_j)$. The latitudes needed to calculate the variable NoSo are derived from CEPII.

$Remote_{it}$ is and indicator of remoteness of country *i* in year *t*, calculated as a GDP weighted average of distance to the countries with which country *i* trades:

$$\mathrm{Re}mote_{it} = \ln\left(\sum_{j}^{T(i)} \frac{GDP_{jt}}{GDP_{Wt} - GDP_{it}} \times Dis\tan ce_{ij}\right)$$, where Distance_{ij} is the distance between *i* and *j*, GDP_{Wt} is the world GDP in year *t*, and *T(i)* is the number of the destination countries of exports from *i*. *T(i)* can vary for each *i*, for instance, when *i* is a composite, the number of destination countries is 95, while when *i* is an individual countries, *T* is 94. Data on trade flows, *ad-valorem* trade protection and GDP in each country are taken from the Global Trade Analysis Project (GTAP) 6.0 and 7.0 databases.

FTA_{ij} is a dummy variable that values 1 when countries *i* and *j* have belonged to the same Free Trade Agreement Area, at least from 2001. FTA includes EU15, Nafta, Mercosur (Southern Cone Common Market), Andean Pact, Caricom (Caribbean Community and Common Market) and CACM (Central American Common Market).

Gdp_{ijt} is the product of GDP in country *i* and country *j* in year *t*, in logs: $\ln(GDP_{it} \times GDP_{jt})$, with GDP measured in million US$ in 2001, and 2004;

$Sqincome_{ijt}$ is the square of difference in per capita income in countries *i* and *j*, in logs: $\ln((GDPpc_{it} - GDPpc_{jt})^2)$ with GDPpc measured in US$ per habitant in 2001, and 2004. The population data needed to calculate per capita income is obtained from "World Bank, Development indicators": http://devdata.worldbank.org/data-query

$Year2004_t$ is a dummy variable that values 1 when the year *t* is 2004;

F_i (F_j) is the fixed effect for exporter (importer) country *i (j)* or a dummy variable that values 1 when the exporter (importer) is *i (j)*.

Bilateral distance (Dist), contiguity (Cont), common official language (Lang) data are taken from CEPII (Centre d'Études Prospectives et d'Informations Internationales), www.cepii.fr.

Data used in the estimation have been re-calculated when necessary to accommodate the GTAP regional aggregation, and re-arranged into datasets where each observation represents a bilateral relationship. The latter leads to databases with 17892 observations for each sector. Bi-lateral data when origin and destination are the same region are removed unless the region is composite.

3.2. Results and discussion

We use a Poisson Maximum Likelihood approach as proposed by Silva and Tenreyro (2006) to estimate the gravity equation. Fixed effects for importer countries are included, while only a sub-set of fixed effects for exporters are included to avoid the multi-collinearity. Table 1 provides a description of the 20 agro-food sectors, which have been analysed in this paper. Results of the estimation for each sector are shown in Tables 2 and 3 for agricultural and food products, respectively.

The Vuong test (1989) has been applied to compare between Poisson and ZIP models, as well as between NB and ZIB. Results indicate that in many sectors zero-inflated models are favoured against the non-inflated versions.

A further issue that of over-dispersion, is confirmed in most of the sectors by two tests (LR test and the individual significance test of α, the over-dispersion parameter). The presence of over-dispersion favours the selection of the NB against the Poisson model, and the ZIB versus ZIP model.

Unfortunately, neither of the two tests above (i.e., the Vuong test and the likelihood ratio test of over-dispersion) can be used to compare Poisson and ZIB, as well as with NB and ZIP. Additionally, some selected models according to the Vuong test and/or the over-dispersion LR test present convergence problems. When this occurs, the model with a lower Akaike Information Criteria (AIC), and also an improvement in the individual significance of the tariff and the subsidy coefficients constitutes the best choice. For each sector, the final choice of model has been indicated by a superscript in the first column of Tables 2 and 3.

Focusing first on trade barriers, we observe that in all regressions, *ad-valorem* tariffs have the expected negative sign, although in two agricultural sectors ('osd' and 'c_b') tariffs are not significant. Our results represent an improvement on Kuiper and van Tongeren (2006) and on Philippidis *et al.* (2007), who find various examples of positive tariff parameter values across sectors. As noted previously, the authors suggest that this may be linked to the inadequate treatment of non-tariff barriers (NTBs) in the econometric specification.

Focusing on export subsidies, a positive and significant coefficient is found in eight agro-food sectors, whilst an unexpected negative but non-significant effect is found in 'cmt', and 'b_t'.

Moving to non-tariff barriers to trade, distance has a negative and significant impact on trade in every sector. The median parameter estimate is -1.133, with a minimum impact of -0.494 in 'pfb' and a maximum of -1.81 in 'gro'. Thus, in 50% of the sectors, trade increases by more than 1.133% when distance drops by 1%. The most sensitive trade changes to distance, with parameters greater than 1.50 in absolute value, occur mainly in agricultural sectors ('ctl', 'c_b', and 'wht'), but also in 'mil'. Importantly, the range of distance estimates lies within the boundaries provided by the empirical literature. However, when the distance is measured along the axis North-South, a positive and significant effect is found in the majority of sectors, although the effect is of little magnitude (0.042 in average). On the other hand, in four sectors a negative and significant impact is found contrary to expectations, intensifying the already negative distance effect ('osd', 'oap', 'omt' and 'pcr').

Contiguity is also found to favour significantly trade in almost every sector. In half of the sectors, the fact of sharing a border increases bilateral trade by more than 1.674 times (exp (0.515)) with respect to those countries which are not contiguous. Language facilitates significantly trade only in 7 sectors (osd, omt, vol, mil, sgr, ofd and b_t), with an average impact of 0.118: countries with the same official language trade 1.125 times more than those countries with no common language.

Colonial links are also found to encourage greater bilateral trade in 13 sectors, with an average impact of 0.411.

In many sectors remoteness plays a significant role in explaining bilateral trade. As expected, the more remote a country, the greater is the likelihood that it trades with a specific partner, *ceteris paribus*. The median impact is 2.723, so in 50% of the sectors, a 1% increase in remoteness increases bilateral trade by more than 2.723%. Nevertheless, a significant negative impact is also found in 3 sectors ('rmk', 'wol' and 'pcr').

Being member of a regional trade agreement, increases significantly bilateral trade in 10 sectors. The impact is of a larger magnitude in the agricultural sectors ('wht', 'pdr', and 'gro' by this order). The positive impacts range between 0.144 (vol) and 1.514 (wht). Countries within the same free trade area trade between 1.15 (exp(0.144)) and 4.54 (exp(1.514)) times as much with each other as countries which are not members of the same free trade areas. However, we also find negative significant effects in 3 sectors ('ctl', 'omt' and 'sgr').

Countries with more similar *per capita* incomes trade more in 6 sectors ('wht', 'ctl', 'wol', 'mil', 'pcr', and 'b_t'), which implies that a significant Linder effect is found in these sectors. On the other hand, in other 5 sectors ('pdr', 'pfb', 'oap', 'sgr' and 'ofd'), the bigger is the difference between per capita income, the larger is the bilateral trade volume after controlling for other influences.

Table 1. Description of the 20 sectors and its correspondence with the sector code

Sector code	Name of the sector	Definition
pdr	Paddy rice	Husked and not husked
wht	Wheat	Soft and durum wheat
gro	Other cereal grains	Rye, sorghum, barley, oats, maize, millet and other cereals
v_f	Vegetables, fruits and nuts	All vegetables, fruits and nuts
osd	Oilseeds	Oilseeds and oleaginous fruits
c_b	Raw sugar	Sugar cane and sugar beet
pfb	Plant based fibers	Raw vegetable materials used in textiles; seeds, live plants
ocr	Other Crops	Other crops
ctl	Cattle, sheep, goats and horses	Live bovine cattle, sheep and goats for fattening, horses, asses, mules
oap	Other animal products	Live swine and poultry for fattening, other animals; eggs, honey, snails and frogs legs
rmk	Raw milk	Dairy and other cows
wol	Wool and silk-worm cocoons	Wool and silk worm cocoons
cmt	Meat of cattle, sheep, goats and horses	Meat of cattle, sheep, goats and horses
omt	Other meat products	Other meat products including edible offals and animal oils and fats
vol	Vegetable oils and fats	Oils of: Coconuts, cottonseeds, groundnuts, oilseeds, olives, palmkernels, rice brans, rape and mustard, soyabeans, sunflower seeds; and fats
mil	Dairy products	All dairy products
pcr	Processed rice	Milled rice
sgr	Sugar	Refined sugar, sweeteners
ofd	Other food products	Prepared and preserved sea food products, vegetables and fruits, bakery and confectionary products, pastas and flours
b_t	Beverages and tobacco products	Cigarettes, cigars etc., wines and spirits, beer

Table 2. Estimated parameters of the gravity equation on agricultural products

Sector code		Mt	Xs	Dist	Contig	Lang	Colony	NoSo	Remote	FTA	gdp	Sqincome	Year 2004	Constant	LR	AIC	Dispersion Parameter Coef.	Confidence Interval Lower	Confidence Interval Upper	Vuong Test
pdr[3]	coef	-0.201	.	-0.553	0.340	-0.023	0.479	-0.002	5.867	1.072	1	0.044	-0.439	-71.159	1987.042	0.418	4.451	1.465	13.515	0.630
	p-valor	0.047	.	0.000	0.009	0.785	0.000	0.962	0.001	0.000	.	0.001	0.000	0.000	0.000					0.264
wht[1]	coef	-0.855	3.788	-1.568	0.109	-0.058	0.931	0.018	3.553	1.514	1	-0.015	-0.248	-43.101	34126.926	2.774	.	.	.	2.280
	p-valor	0.000	0.000	0.000	0.000	0.010	0.000	0.012	0.000	0.000	.	0.000	0.000	0.000	0.000					0.011
gro[4]	coef	-0.810	2.755	-1.810	0.192	0.245	0.343	0.058	3.531	0.895	1	0.002	-0.118	-47.076	Na.	0.050	0.838	0.707	0.994	N.a.
	p-valor	0.013	0.004	0.000	0.250	0.111	0.186	0.160	0.185	0.000	.	0.904	0.356	0.044	Na.					N.a.
v_f[3]	coef	-0.949	2.675	-1.056	0.948	-0.040	0.391	0.130	3.001	0.310	1	0.004	-0.195	-38.361	1.398	0.046	0.485	0.266	0.882	0.210
	p-valor	0.000	0.014	0.000	0.000	0.588	0.000	0.000	0.013	0.011	.	0.651	0.001	0.000	0.237					0.419
osd[4]	coef	-0.082	.	-1.316	0.722	0.674	0.252	-0.196	7.218	0.207	1	0.041	-0.214	-80.992	N.a	0.057	0.799	0.661	0.965	N.a.
	p-valor	0.887	.	0.000	0.000	0.000	0.307	0.000	0.029	0.418	.	0.109	0.176	0.005	N.a					N.a.
c_b[3]	coef	-1.011	.	-1.576	2.164	0.593	0.550	0.532	-9.995	-0.053	1	0.107	-0.716	48.516	23.602	0.042	0.571	0.247	1.323	-0.500
	p-valor	0.126	.	0.000	0.033	0.118	0.446	0.043	0.441	0.962	.	0.110	0.210	.	0.000					0.691
pfb[3]	coef	-4.937	0.662	-0.494	0.486	-0.126	0.195	-0.010	5.009	0.430	1	0.039	-0.334	-75.040	11333.866	1.201	9.143	6.553	12.756	0.700
	p-valor	0.000	0.850	0.000	0.000	0.085	0.095	0.722	0.000	0.003	.	0.000	0.000	0.000	0.000					0.243
ocr[1]	coef	-0.341	.	-0.928	0.259	-0.125	0.140	0.083	2.483	-0.153	1	0.005	-0.589	-35.950	14.106	4.763	.	.	.	2.060
	p-valor	0.050	.	0.000	0.003	0.139	0.204	0.002	0.092	0.103	.	0.549	0.000	0.005	0.000					0.020
ctl[4]	coef	-1.010	2.403	-1.644	1.249	-0.201	0.621	0.039	2.379	-0.346	1	-0.058	-0.341	-34.002	11987.406	0.559	1.436	1.347	1.531	-2.780
	p-valor	0.025	0.044	0.000	0.000	0.073	0.001	0.332	0.275	0.029	.	0.000	0.001	0.075	0.000					0.997

Table 2 (cont.). Estimated parameters of the gravity equation on agricultural products

Sector code		Mt	Xs	Dist	Contig	Lang	Colony	NoSo	Remote	FTA	gdp	Sqincome	Year 2004	Constant	LR	AIC	Dispersion Parameter			Vuong Test
																	Coef.	Confidence Interval Lower	Confidence Interval Upper	
oap[3]	coef	-1.068	.	-0.528	0.882	0.002	0.223	-0.068	2.213	0.676	1	0.051	-0.552	-36.158	14475.290	1.804	4.095	3.259	5.145	-2.490
	p-valor	0.000	.	0.000	0.000	0.968	0.004	0.019	0.081	0.000	.	0.000	0.000	0.001	0.000					0.994
rmk[2]	coef	.	0.438	-0.497	0.441	0.009	0.175	0.015	-16.857	0.520	1	0.001	-0.963	90.021	0.001	0.061	.	.	.	-0.600
	p-valor	.	0.352	0.004	0.107	0.973	0.537	0.835	0.049	0.189	.	0.992	0.045	.	1.000					0.727
wol[2]	coef	-8.859	.	-0.716	0.374	0.096	0.447	0.128	-12.139	0.196	1	-0.026	-1.549	86.943	0.000	0.105	.	.	.	N.a.
	p-valor	0.000	.	0.000	0.027	0.396	0.051	0.001	0.000	0.344	.	0.052	0.000	0.000	1.000					N.a.

[1]ZIP model, [2]Poisson model, [3]NB model, [4]ZINB model. Note: Results for the country-specific dummies are not reported. P-values lower than 0.05 implies significant coefficient at 5%, except for sector c_b (significance at 15%). "N.a" indicates that the corresponding value is not available. The inflated constant (τ) was significant at 1% in all zero-inflated models, results are not shown in this table for space reasons. The theory predicts that the parameter of the log-product of GDP (or each of the individual log-GDPs) is one. In our empirical application, we include the product of GDPs and impose the restriction that the parameter is one.

Table 3. Estimated parameters of the gravity equation on food products

Sector code		Mt	Xs	Dist	Contig	Lang	Colony	NoSo	Remote	FTA	gdp	Sqincome	Year 2004	Constant	LR	AIC	Dispersion Parameter Coef.	Confidence Interval Lower	Confidence Interval Upper	Vuong Test
cmt[4]	coef	-0.415	-0.016	-1.265	0.665	0.060	0.802	0.026	2.092	0.752	1	0.007	-0.360	-32.895	46177.830	1.197	1.457	1.388	1.530	1.220
	p-valor	0.013	0.946	0.000	0.000	0.516	0.000	0.397	0.173	0.000	.	0.563	0.000	0.015	0.000					0.110
omt[4]	coef	-1.365	2.098	-1.252	2.141	0.259	0.557	-0.038	3.720	-0.303	1	-0.009	-0.380	-31.529	N.a	23.176	2.446	2.404	2.489	N.a.
	p-valor	0.000	0.122	0.000	0.000	0.000	0.000	0.010	0.000	0.002	.	0.087	0.000	0.000	N.a					N.a.
vol[3]	coef	-0.718	12.438	-0.579	0.496	0.174	0.190	0.012	6.900	0.144	1	0.007	-0.078	-75.774	71683.742	1.948	23.036	19.455	27.276	2.810
	p-valor	0.000	0.025	0.000	0.000	0.001	0.002	0.438	0.000	0.088	.	0.133	0.037	0.000	0.000					0.003
mil[4]	coef	-0.707	1.025	-1.531	1.266	0.413	0.404	-0.031	5.123	0.429	1	-0.032	0.097	-41.449	N.a	23.299	2.226	2.199	2.253	2.320
	p-valor	0.000	0.000	0.000	0.000	0.000	0.001	0.272	0.000	0.002	.	0.000	0.071	0.000	N.a					0.010
pcr[1]	coef	-1.171	0.275	-1.209	0.211	-0.243	0.432	-0.176	-2.371	0.739	1	-0.048	-0.256	15.870	N.a	1.090	.	.	.	23.630
	p-valor	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000	0.000	N.a					0.000
sgr[3]	coef	-0.521	0.533	-1.305	0.533	0.260	0.156	0.126	1.466	-0.407	1	0.027	-0.473	-39.592	37917.820	0.941	26.943	24.530	29.593	-0.530
	p-valor	0.000	0.154	0.000	0.000	0.000	0.143	0.000	0.309	0.006	.	0.010	0.000	0.951	0.000					0.702
ofd[3]	coef	-0.736	2.725	-0.724	0.627	0.117	0.310	0.027	1.315	0.695	1	0.020	-0.305	-25.461	156273.820	4.290	19.162	18.459	19.892	0.500
	p-valor	0.000	0.016	0.000	0.000	0.000	0.000	0.000	0.002	0.000	.	0.000	0.000	0.000	0.000					0.308
b_t[2]	coef	-0.825	-9.560	-1.012	0.234	0.281	0.621	0.173	2.963	0.187	1	-0.035	-0.241	-40.232	101452.110	7.858	.	.	.	-0.740
	p-valor	0.010	0.567	0.000	0.095	0.002	0.000	0.000	0.158	0.225	.	0.008	0.012	0.029	0.000					0.769

[1]ZIP model, [2]Poisson model, [3]NB model, [4]ZINB model. Note: Results for the country-specific dummies are not reported. P-values lower than 0.05 implies significant coefficient at 5%, except for sectors omt, and sgr (significance at 15%). "N.a" indicates that the corresponding value is not available. The inflated constant (τ) was significant at 1% in all zero-inflated models, results are not shown in this table for space reasons. The theory predicts that the parameter of the log-product of GDP (or each of the individual log-GDPs) is one. In our empirical application, we include the product of GDPs and impose the restriction that the parameter is one.

Conclusions

A theoretically consistent gravity equation has been estimated for 20 sectors, using bilateral trade and policy variable data observations from 2001 and 2004, supported by further economic, geographical and cultural indicators across 95 regions. The standard OLS approach has been replaced by the Poisson Maximum Likelihood Estimator and some variants of the Poisson model (Negative Binomial, Zero Inflated Poisson, and Zero Inflated Negative Binomial), which provide consistent estimates, whilst allowing for the inclusion of zero-trade values.

The sectoral approach in the estimation of gravity equations provides heterogeneous results in terms of the influence of specific explanatory variables on trade, which makes it difficult to draw clear conclusions. It is perhaps for this reason that most gravity applications are restricted to an analysis and discussion of 'total trade' flows only. In general, the estimated coefficients match theoretical expectations and their magnitude is in accordance with other empirical studies. In this sense, import tariffs constrain trade significantly in almost all sectors, while export subsidies favour trade in eight agro-food sectors. Notwithstanding, as a policy tool a greater number of insignificant coefficients may indicate its declining usage in EU trade policy. Distance plays an important role impeding trade, (especially in wheat and other cereal grains, cattle, raw sugar and dairy products), while this effect is partly mitigated by North-South hemisphere and remoteness indicators. Contiguity and a colonial relationship contribute significantly to increase trade, while easier communications induced by a common language only favour trade in some sectors. One of the most influential variables explaining bilateral trade is in those cases where both partners are members of a regional free trade agreement; this explanatory variable has a particularly notable impact in wheat and other cereal grains, and paddy rice. In a significant group of sectors countries with more similar per capita incomes trade more prolifically, supporting the Linder hypothesis.

The use of panel data with disaggregated information for 20 sectors can be viewed as a key novelty of the research, and may be especially useful for policy makers. Panel data with fixed effects are employed to improve the estimation of the parameters of interest in the gravity equation. The approach developed in this research mitigates the problem of multi-collinearity, and has the advantage of including a detailed sectoral disaggregation. Although the Poisson model and its variants (Negative Binomial, Zero Inflated Poisson and Zero Inflated negative Binomial) have been proved to have better properties than OLS, model misspecification issues may still persist leading to inconsistent estimates. In particular, endogeneity of trade barriers is still present, while omitted variables, such as the quantifiable existence of NTBs, may be leading to biased parameter estimates.

Acknowledgments

The authors gratefully acknowledge financial support from the Department of Environment, Food, Agriculture and Rural Affairs (DEFRA) of the United Kingdom. The authors would also like to thank Patricia Sisó in helping with the collection of data.

References

Anderson J.E. (1979). Theoretical foundation of the gravity equation. *American Economic Review 69(1)*, 106-116.

Armington P. (1969). A theory of demand for products distinguished by place of production. IMF Staff Papers 16, 159-176.

Bergstrand J.H. (1990). The Heckscher-Ohlin-Samuelson model, the Linder hypothesis and the determinants of bilateral intra-industry trade. *The Economic Journal 100*, 1216-1229.

Burger M., van Oort F. and Linders G.-J. (2009). On the specification of the gravity model of trade: zeros, excess zeros and zero-inflated estimation. *Spatial Economic Analysis 4:2*, 167-190.

Cameron, T. A. and Englin, J. (1997). Respondent experience and contingent valuation of environmental goods. *Journal of Environmental Economics and Management* 33: 296-313.

Cameron A.C. and Trivedi P.K. (1998). *Regression Analysis of Count Data*. Cambridge University Press.

Deardorff A. (1998). "Determinants of bilateral trade: Does gravity work in a neoclassical world?" in Frankel J. (Ed): *The regionalization of the world economy*. The University of Chicago Press.

Durham, C. A., Pardoe, I. and Vega, E. (2004). A methodology for evaluating how product characteristics impact choice in retail settings with many zero observations: An application to restaurant wine purchase. *Journal of Agriculture and Resource Economics* 29: 112-131.

Edmeades, S. and Smale, M. (2006). A trait-based model of the potential demand for a genetically engineered food crop in a developing economy. *Agricultural Economics* 35: 351-361.

Feenstra, R.C. (2004). *Advanced International Trade: Theory and Evidence*, Princeton University Press.

Greene, W. (2002). LIMDEP version 8.0: Econometric modelling guide, Plainview, New York.

Kuiper M. and van Tongeren F. (2006). "Using gravity to move Armington. An empirical approach to the small initial trade share problem in general equilibrium models". Ninth Annual Conference on global Economic Analysis, Addis Ababa (Ethiopia).

Lambert, D. (1992). Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics* 34: 1-14.

Martin, T. G., Wintle, B. A, Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J. and Possingham, H. P. (2005). Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters* 8: 1235-1246.

Philippidis, G., Renwick, A. and Sanjuán, A.I. (2007). *Defra-TAP Tailoring the GTAP Model for Defra Project EPES 0607/08,* Technical Book for the Department of Environment, Food and Rural Affairs (DEFRA), UK Government.

Shankar, V., Milton, J. and Mannering, F. (1997). Modelling accident frequencies as zero-altered probability processes: An empirical inquiry. *Accident Analysis and Prevention* 29: 829-837.

Silva J.M.C.S. and Tenreyro S. (2006). The Log of Gravity. *The Review of Economics and Statistics 88(4)*, 641-658

Tinbergen J. (1962). *Shaping the world economy*. New York: The Twentieth Century Fund.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57: 307-333.

Westerlund J. and Wilhelmsson F. (2006). Estimating the gravity model without gravity using panel data. Mimeo School of Economics and Management, Lund University. http://www.nek.lu.se/nekjwe/research.htm

Wang, C., Wei, Y. and Liu, X. (2010). Determinants of bilateral trade flows in OECD countries: Evidence from gravity panel data models. *The Worl Economy 33(7), 894-915.*