European Commission

JRC SCIENCE AND POLICY REPORT

Gravity estimations to correct the 'small shares stay small' bias in economic models.

The example of Mercosur and EU agri-food trade

Ana I. Sanjuán López
Helena Resano Ezcaray

2015

European Commission
Joint Research Centre
Institute for Institute for Prospective Technological Studies

Contact information
Robert M'barek
Address: Joint Research Centre, xxx *(relevant complete address)*
E-mail: Robert.M'BAREK@ec.europa.eu
Tel.: + 34 954 48 8489

https://ec.europa.eu/jrc
https://ec.europa.eu/jrc/en/institutes/ipts

JRC96089

EUR 27264

ISBN 978-92-79-48391-2 (PDF)

ISSN 1831-9424 (online)

doi:10.2791/531517

Luxembourg: Publications Office of the European Union, 2015

Abstract

The simulation of liberalization trade scenarios in economic models normally understate the export growth for countries with small initial trade shares but which nevertheless could be competitive under a new tariff regime. This downward bias is known as the ´small share stay small´ and it is inherent to the constant elasticity of substitution in the Armington demand specification. In this report, we show how the gravity equation can provide econometric estimates of the tariffs restrictiveness and trade shares after tariff liberalization and how these can be input into a General Equilibrium (CGE) model to remedy said bias. The fusion approach between gravity and CGE that we follow closely in this report was proposed by Kuiper and van Tongeren (2006) and further developed by Philippidis et al.(2014). As an empirical illustration, the method is applied to agro-food trade between EU and Mercosur where a pervasiveness of 'small-share' examples exists.

Acknowledgments

We thank George Philippidis, Robert M'barek, Koen Dillen, Sophie Hélaine, Ben Van Doorslaer and Pierluigi Londero for providing helpful comments. Any errors remain the responsibility of the authors.

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

TABLE OF CONTENTS

List Of Tables

Tables In Appendix

1. INTRODUCTION

The aim of this report is to explain a methodological approach that allows alleviating the 'small shares stay small' problem. The ‘small shares stay small’ problem arise when due to the initial ‘small’ observed import shares of importer countries or regions, a final ‘small’ trade creation effect is obtained even after simulating ambitious liberalisation scenarios implying noticeable tariff barriers removal. This result is consequence of the Armington (1969) specification in the trade demand functions.

The methodology proposed follows the proposal by van Kuiper and van Tongeren (2006) and applied by Philippidis *et al*. (2014), which consists, in a first stage, in calculating predicted import shares after full (or partial) import tariff liberalisation for the bilateral trade route under study through a gravity equation.In a second stage, these new predicted trade shares, in combination with the original market shares embedded in the model dataset, allows to provide a shifter for the demand function in the economic model .In other words, the gravity equation provides the basis for calculating a supplementary import demand shock for trade creation, which when inputted into the General (CGE) or Partial Equilibrium(PE) model, will hasten the trade diversion effects further on other partner countries (given the strength of the Armington elasticities and primary factor resource constraints).

In this report, we apply the Poisson family models to estimate sectoral gravity equations, and as an illustration, we contemplate trade liberalization between the EU and Mercosur. As Philippidis *et al*. (2014) pointed out, EU–Mercosur trade relations may be considered as ‘small share’ examples, as there are some agro-food sectors where high tariffs co-exist with small shares, such as ´red´ and ´white meat´, ‘dairy products’, ‘processed rice’ and ‘sugar’ sectors.

This report could serve as a useful input to deal with the ‘small shares’ issue allowing the inclusion of predicted import shares by gravity equations into the current models used in the AGRILIFE unit in JRC-IPTS (partial and general equilibrium models, in particular, CAPRI and MAGNET).

The report is structured as follows: Section 2 describes the methodology, contemplating both, the gravity model and the implementation of its results into a CGE model; Section 3 describes the data sources and the gravity model specification; Section 4 presents briefly the main results obtained after full tariff elimination in the bilateral trade route under study; and Section 4 presents some final remarks.

2. METHOD

2.1. THE GRAVITY EQUATION

In its simplest form, the gravity model posits that trade between two countries is a positive function of GDP (i.e., 'mass') and a negative function of trade costs (i.e., distance) (Tinbergen, 1962, Pullianen, 1963). Empirical applications have extended this specification to encompass (*inter alia*) preferential trade (Kandogan, 2008; Foster et al., 2011; Hayakawa and Yamashita, 2011), contiguity (Bergstrand, 1985; Thoumi, 1989), common language and/or ex-colonial ties (e.g. Rose and van Wincoop, 2001), or even to cater for the effect of distance along different hemispheres as well as remoteness (Melitz, 2007), and being landlocked (MacPhee *et al.*, 2013). Other developments (Arnon et al., 1996; Hallack, 2006) account for the so called 'Linder' hypothesis (Linder, 1961), which posits that countries with similar *per capita* incomes have a greater tendency to engage in mutual trade. This is seen as a test of the monopolistic intra-industry hypothesis, whilst the polar opposite that differences in per capita incomes (which proxy for differing factor intensities) promote trade can be interpreted as support for the Heckscher-Ohlin (HO) hypothesis.

Along time, the gravity equation has regained credibility as it has been substantiated in economic theory. Anderson and van Wincoop (2003) formalised a paradigm for subsequent econometric gravity work, providing an explicit treatment of prices whilst accommodating the empirical observation of 'cross-hauling' of differentiated products. Expressed as a CES preference function of the form:

$$X_{ij} = \frac{Y_j Y_i}{Y_w} \left(\frac{t_{ij}}{P_j \Pi_i} \right)^{1-\sigma} \qquad (1)$$

where X_{ij} are exports from country *i* to country *j*; Y_i and Y_j represent GDP, Y_w is world GDP, t_{ij} are trade costs $t_{ij} = 1 + \tau_{ij}$ (where τ_{ij} is expressed as 'iceberg cost' [1], and no trade costs imply t_{ij}=1); and σ is the elasticity of substitution between varieties (i.e. countries).
The variables Π_i and P_j are price indices, capturing the relative country's international competitiveness. These terms are denominated as 'multilateral resistance' terms. which are dependent of bilateral trade barriers (t_{ij}), and which reflect how difficult it is for a country

[1] The concept of iceberg cost was developed by Samuelson (1952), who suggested that some fraction of a commodity 'melts' away as a necessary cost of transportation over a unit of distance. This construct is equally applicable to trade costs, which inhibit the effective flow of goods and services from one region to another.

to trade with the rest of the world. Anderson and van Wincoop (2003, 2004) stress the importance of controlling for these multilateral resistance terms arguing that trade between two regions depends on the bilateral barrier between them relative to the average trade barriers that both regions face with all their trading partners. Baier and Bergstrand (2010) simplify the non-linear approximation employed by Anderson and van Wincoop (2003) by proposing exogenous multilateral resistance terms defined in terms of GDPs and trade restriction measures, such as distance and borders. Concurring with previous literature (Feenstsra, 2004; Anderson and van Wincoop, 2004), Baier and Bergstrand (2009) also find the empirical result that unobserved multilateral resistance terms can be proxied with importer and exporter fixed effects.[2]

2.2. POISSON FAMILY MODELS

From an econometric standpoint, earlier studies favoured the use of an Ordinary Least Squares (OLS) log-linear specification. Subsequent literature (Santos Silva and Tenreyro, 2006; 2011) favours the estimation of the gravity model in its theoretical multiplicative form (see Equation 1), as the log-transformation may lead to substantial bias in coefficient estimates in the presence of heteroskedasticity. This leads Santos Silva and Tenreyro (2006; 2011) to recommend the Poisson estimator, which belongs to the category of count models, which besides adequately cater for zero trade value observations.

[2]Moreover, the use of fixed effects is not only important in producing unbiased estimates (e.g., those associated with the border, the Free Trade Area (FTA)), but they also enter into the calculation of trade impacts from the removal of a border or creation of an FTA (Anderson and van Wincoop, 2003; Baier and Bergstrand, 2009), although the associated AVE is still derived from the coefficient of interest (Anderson and van Wincoop, 2003, p.20).

Trade observations are not pure count, but rather non-negative continuous data. Notwithstanding, the Poisson Maximum Likelihood estimator still provides consistent estimates (Woolridge, 2002), in which case it is referred to more precisely as the Poisson Pseudo-Maximum Likelihood (PML) estimator (Gourieoroux *et al.*, 1984). The model assumes that the observed volume of trade between countries *i* and *j*, X_{ij} follows a Poisson distribution with a conditional mean (μ_{ij}) which is an exponential function of the explanatory variables $\boldsymbol{z}$: $\mu_{ij} = exp(\boldsymbol{\beta}' z_{ij})$[3].

The Poisson model implies equi-dispersion (i.e. the conditional variance equals the conditional mean). However, it is usual to find that the variance exceeds the mean (i.e. over-dispersion). When this occurs, the robust variance estimator may be used to avoid the presence of largely deflated standard errors and consequently, largely inflated t-statistics (Cameron and Trivedi, 2005; 2010). Furthermore, in the presence of over-dispersion, the Negative Binomial (NB) model may be more appropriate.
Two variants of the NB exist (NB1 and NB2). Both consider the same specification for the mean: $\mu_{rs} = \exp(\beta' Z_{rs})$. The difference lies on the variance: NB1 considers the variance as a linear function of the mean, $Var[X_{rs}] = \mu_{rs} + \alpha\mu_{rs}$, where α is the dispersion parameter; while NB2 (the most frequently used because it can be applied to many empirical situations), the variance is a quadratic function of the mean: $Var[X_{rs}] = \mu_{rs} + \alpha\mu_{rs}^2$. Note also that the Poisson is a particular case of NB (where α=0) (Cameron and Trivedi, 2005; 2010).

Summing up, in presence of over-dispersion, both Poisson with a robust estimation of the standard errors and the Negative Binomial (NB2) where the variance is a quadratic function of the mean are suitable (Cameron and Trivedi, 2010), and accordingly, in the empirical application, the Negative Binomial model refers in particular to NB2.

2.3. LIBERALIZATION SIMULATION

Once the gravity equation is estimated, the estimated parameters ($\hat{\beta}$) are saved and the fitted values for each observation ($\hat{X}_{rs}$) are calculated (Cameron and Trivedi, 2010, p.337):

$$\hat{X}_{rs} = \exp(\hat{\beta}' Z_{rs}) \quad (2)$$

[3] See Cameron and Trivedi (1998) for a detailed discussion of count models.

where Z_{rs} are the explanatory variables. Then, the residuals are computed as:

$$\mathrm{res}_{\mathrm{rs}} = \mathrm{X}_{\mathrm{rs}} - \hat{\mathrm{X}}_{\mathrm{rs}} \tag{3}$$

where X_{rs} are the observed values for each observation.

The simulated trade flow after a change in tariffs in the specific bilateral route (i.e. Mercosur to EU, EU to Mercosur) is then computed. To do that, first a new value for the import tariffs in Z_{rs} needs to be assigned (e.g. 0 and 0.5 multiplied by the initial tariff values for a total elimination and a 50% cut in tariffs, respectively) leading to a new matrix Z_{rs}^{*}, and a new set of predicted trade flows:

$$\hat{\mathrm{X}}_{\mathrm{rs}}^{*} = \hat{\beta}' \mathrm{Z}_{\mathrm{rs}}^{*} \tag{4}$$

Instead of directly using these trade flows, Kuiper and Van Tongeren (2006) adjust them by the residuals obtained in (3) relative to the fitted value:

$$\hat{X}_{rs}^{**} = \hat{X}_{rs}^{*}\left(1 + \frac{res_{rs}}{\hat{X}_{rs}}\right) \tag{5}$$

The authors argue that if residuals are ignored, "countries with identical values for the explanatory variables end up with identical fitted trade flows, which may result in large shifts in trade flows with a negligible change in tariffs" (pp. 8). In other words, the final prediction is re-scaled according to the error committed in the estimation.

In this report, we simulate total (100%) removal of bilateral tariffs between EU and Mercosur. Observed and predicted imports (after adjustment (5)) are calculated adding up the individual observations for three alternative origins and destination regions: EU, Mercosur and Rest of the World (RoW), and are called *imptot* and *pred2*, respectively. Based on observed and predicted imports, corresponding trade shares from each origin region *k* = EU, Mercosur and Rest of the World (RoW), to each destination *l* = EU, Mercosur and RoW, ara calculated as:

Observed Trade Share from region *k* to region *l*: $$\mathrm{MSHRS}_{\mathrm{kl}} = \frac{\mathrm{Imptot}_{\mathrm{kl}}}{\sum_{\mathrm{k}=1}^{3} \mathrm{Imptot}_{\mathrm{kl}}} \tag{6}$$

Predicted Trade Share from region k to region l: $$GSHRS_{kl} = \frac{\mathrm{Pred2}_{\mathrm{kl}}}{\sum_{\mathrm{k}=1}^{3} \mathrm{Pred2}_{\mathrm{kl}}} \tag{7}$$

Note however that the tariffs reduction is only simulated for the bidirectional route EU-Mercosur. Therefore, predicted values of trade when the origin or destination region is RoW, do not change (i.e. the gravity model only simulates trade creation but not trade diversion). What can change, however, is the trade shares, in the routes with RoW.

2.4. CGE AND GRAVITY IMPLEMENTATION

The use of the predicted gravity trade shares to alleviate the 'small shares stay small' problem is illustrated in the framework of the GTAP General Equilibrium Model. What follows is a transcription from Philippidis *et al.* (2014, pp.29).

A common representation in the standard GTAP model is the usage of linear (percentage change) behavioural equations, where constant elasticity of substitution (CES) Armington (1969) import demands for commodity 'i', from origin 'r' to destination 's' can be expressed as:

$$m_{i,r,s} = -a_{i,r,s} + mc_{i,s} - \sigma_i^M \left(p_{i,r,s} - a_{i,r,s} - pc_{i,s} \right) \quad (8)$$

$$pc_{i,s} = \sum_r MSHRS_{i,r,s} \left[p_{i,r,s} - a_{i,r,s} \right] \quad (9)$$

where $m_{i,r,s}$ represents bilateral import demand; $p_{i,r,s}$ is the post-tariff market price in 's' on imports from origin 'r'; and σ_i^M is the elasticity of substitution of commodity 'i' between alternative origin routes. The composite import price index ($pc_{i,s}$) in region 's' is calculated as a weighted import value share (MSHRS) of individual import prices from origin 'r' (equation 9), whilst $mc_{i,s}$ is an equivalent aggregate import quantity index in region 's'. In linear form, the small share problem manifests itself by the size of the coefficient 'MSHRS' in equation (9). If the GTAP benchmark data import share (MSHRS) is 'small' on a particular route of origin 'r', significant (tariff induced) falls in 'p', coupled with a large trade elasticity (σ_i^M), will still only result in negligible bilateral import rises.[4]

When implementing 'post-liberalisation' gravity predictions (GSHRS) into the GTAP model, it is not practical to directly substitute into the GTAP benchmark data since this would disrupt

[4] Assume region 's' has two import shares: $MSHRS_{i,1,s} = 0.99$ and $MSHRS_{i,2,s} = 0.01$, which implies that $pc_{i,s} \approx p_{i,1,s}$ and $mc_{i,s} \approx m_{i,1,s}$. Even if $p_{i,2,s}$ fell by 50%, according to equation (9) the impact on $pc_{i,s}$ would be negligible, and consequently the 'trade expansion' effect (i.e., rise in $mc_{i,s}$), in the Armington function would also be very minor. Given that $m_{i,1,s} \approx mc_{i,s}$ and the latter only increases very slightly, then any increase in favour of $m_{i,2,s}$ (i.e., substitution effect) due to the fall in the price from export region 2 ($p_{i,2,s} < 0$) is also small.

the internal consistency of the GTAP database. Kuiper and van Tongeren (2006) view the necessary change in import composition as akin to an adjustment in 'import technology'. More specifically, an exogenous Hicks neutral technological preference shifter for each bilateral route ($a_{i,r,s}$) displaces the Armington import demand curve to mimic the composition of import trade shares predicted by the gravity model (see equations 8 and 9). Thus, a positive shock to the technological preference shifter reduces the 'effective' bilateral import price (i.e., $p_{i,r,s} - a_{i,r,s}$) and increases the 'effective' quantity imported (i.e., $m_{i,r,s} + a_{i,r,s}$). To determine the magnitude of these shocks, it is assumed that the percentage change in the effective import price is based on the following identity (Kuiper and van Tongeren, 2006):

$$MSHRS_{i,r,s}\left[p_{i,r,s} - a_{i,r,s}\right] = GSHRS_{i,r,s}\, p_{i,r,s} \qquad (10)$$

where GSHRS represent gravity based predictions of import shares after (in the context of this study) liberalisation of tariffs between the EU and Mercosur (i.e. the $GSHR_{kl}$ in Equation (7)). Rearranging (10) in terms of the preference shifter ($a_{i,r,s}$) gives:

$$a_{i,r,s} = p_{i,r,s} - \frac{GSHRS_{i,r,s}}{MSHRS_{i,r,s}}\, p_{i,r,s} \qquad (11)$$

Employing a first order linear price linkage assumption, (*ceteris paribus*) a reduction in an applied tariff of 5%, reduces the market price of imports in region 's' by 5%, whilst the required shock on the preference shifter is positive when the gravity-to-GTAP bilateral import share ratio in equation (11) is greater than one. An elegant feature of equation (11) is that if the gravity model predicts that non-economic factors are the main (only) trade restricting factor, then post trade liberalisation values of GSHRS will remain similar (or equal) to MSHRS. Thus, irrespective of the magnitude of the tariff induced change in the import price, the preference shifter shock in those cases will be small (or zero).

3. DATA AND MODEL SPECIFICATION

3.1. DATA

One principal aim of this study is to provide a methodology for correcting the ´small share´ bias compatible for modelling databases, such as GTAP. Based on the GTAP sector concordance, 10 agri-food sectors are considered (see Table 1) selected according to their weight in total EU agri-food trade.

Bilateral trade and applied *ad-valorem* tariff equivalents for years 2001, 2004, 2007 and 2011, 78 countries (see Table A.1 in the Appendix)are taken from releases 6 (Dimaranan, 2006), 7 (Narayanan and Walmsley, 2008), 8 (Narayanan et al., 2012) and the pre-release 1 of release 9 of the GTAP database. The final dataset constitutes an update of the one used by Philippidis *et al*. (2014), covering more within country variability with the inclusion of two more recent years of information (2007 and 2011).

TABLE 1 DESCRIPTION OF THE AGRI-FOOD SECTORS AND ACCOMPANYING CODES

Sector code	Name of the sector	Definition
wht	Wheat	Soft and durum wheat
v_f	Vegetables, fruits and nuts	All vegetables, fruits and nuts
cmt	Meat of cattle	Meat of cattle, sheep, goats and horses
omt	Other meat products	Meat of swine, poultry, edible offal
vol	Vegetable oils and fats	Oils of: Coconuts, cottonseeds, groundnuts, oilseeds, olives, palm kernels, rice brans, rape and mustard, soyabeans, sunflower seeds; and fats
mil	Dairy products	All dairy products
pcr	Processed rice	Milled rice
sgr	Sugar	Raw and 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, wines and spirits, beer

Other important secondary data sources needed to complete the database used in the gravity equation estimation are the World Bank (2014) for population and GDP; and Centre d'Études Prospectives et d'Informations Internationales (CEPII, 2014) (Mayer and Zignago, 2011) for cultural and geographical distance.

3.2. MODEL SPECIFICATION

In each of the ten sector regressions, the dependent variable of bilateral imports is determined by a number of explanatory variables described in Table 2. The final gravity specification is presented in equation (1). The sub-index *r* and *s* refer to the exporter and importer countries, respectively, whilst *t* refers to the year:

$$E[X_{rst}] = \mu_{rst} = \exp\begin{pmatrix} \beta_0 + \beta_1 Mt_{rst} + \beta_2 Dist_{rs} + \beta_3 DistInt_s + \beta_4 NoSo_{rs} + \beta_5 Lock_r + \beta_6 \operatorname{Re}mote_{rt} \\ +\beta_7 Contig_{rs} + \beta_8 Lang_{rs} + \beta_9 Col_{rs} + \beta_{10} PTA_{rs} + \beta_{11} SqIncome_{rst} + \beta_{12} Gdp_{rst} \\ + \sum_{t=2004}^{2011} \theta_t Y_t + \sum_{r=2}^{N} \theta_r F_r + \sum_{s=2}^{N} \theta_s F_s \end{pmatrix} \quad (12)$$

TABLE 2VARIABLE DESCRIPTIONS IN THE GRAVITY EQUATION

Variable	**Description**
X_{ijt}	Value of imports into country *j* from country *i* at current prices in year t (million US$)
Mt_{ijt}	Power of the import tariff rate ($AdvRate_{ijt}$) applied by importer *j* on imports from *i* in year *t*, measured in *ad-valorem* equivalents, in logs: $Mt_{ijt} = ln\left(1 + \frac{AdvRate_{ijt}}{100}\right)$
$Dist_{ij}$	Weighted bilateral distance by internal population, between the main cities of country *i* and *j,* in logs
$DistInt_j$	Internal distance of country *j*, calculated as $0.67\sqrt{area_j/\pi}$, in logs
$NoSo_{ij}$	Difference in latitudes between countries *i* and *j*, in logs: $ln(abs(latitude_i - latitude_j))$
$Remote_{it}$	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: $Remote_{it} = ln\left(\sum_{j}^{T(i)} \frac{GDP_{jt}}{GDP_{Wt} - GDP_{it}} \times Dist_{ij}\right)$ where $Dist_{ij}$ is the distance between *i* and *j* (defined as above), $GDPW_t$ is the world GDP in year *t*, and T(i) is the number of the destination countries of exports from i.
$Lock_i$	Dummy variable that values 1 when country *i* is landlocked, and 0 otherwise.
$Contig_{ij}$	Dummy variable that values 1 when countries *i* and *j* share a border, and 0 otherwise
$Lang_{ij}$	Dummy variable that values 1 when countries *i* and *j* share the same language (at least 9% of the population speaks it), and 0 otherwise
Col_{ij}	Dummy variable that values 1 when countries *i* and *j* have or have had a colonial linkage
PTA_{ijt}	Dummy variable that values 1 when countries *i* and *j* belong to the same Preferential Trade Area. PTA includes EU, EFTA, NAFTA, CEFTA (Central European Free Trade Agreement), Mercosur (Southern Cone Common Market), Andean Community, Caricom (Caribbean Community and Common Market), CACM (Central American Common Market) and ASEAN (Asean Free Trade Area).
$Sqincome_{ijt}$	Square of the 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 nominal terms)
Gdp_{ijt}	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 nominal terms)
Y_t	Fixed effect for each year t (t= 2002 to 2011), i.e., a dummy variable that values 1 when the year is *t*, and 0 otherwise.
F_i (F_j)	Fixed effects for exporter (importer) country *i* (*j*). $F_i(F_j)$ are dummy variables, that value 1 when the exporter (importer) is *i* (*j*), and 0 otherwise

Thus, the gravity specification includes geographical explanatory factors (i.e., weighted bilateral distance between partners[5] (Dist) and internal distance (DistInt), North-South hemisphere distance (NoSo) and remoteness (Remote), being landlocked (Lock), border sharing dummy (Contig)), historical and cultural linkages (i.e., common language (Lang) and colonial ties (Col) dummies). Additionally, membership of a preferential trade agreement (PTA), the squared difference of per capita GDPs (SqIncome) (Linder hypothesis) and the product of GDPs (Gdp) are incorporated[6]. Finally, bilateral import tariffs (Mt) are inserted into the gravity regression. The gravity equation includes fixed effects for both time and country (exporter and importer).[7] The country fixed effects proxy the unobserved theoretical multilateral resistance terms posed by Anderson and van Wincoop (2003, 2004), while both, country and year fixed effects control for correlation between omitted and observed variables.

[5] Weighted bilateral distance is recommended by Mayer and Zignago (2011) and applied, among others, by Engelbert et al. (2014).

[6]The 'Gdp' coefficient is restricted to a value of 1 according to the theoretical model in equation (1) (Anderson and van Wincoop, 2003; Baier and Bergstrand, 2009). Moreover, in comparison to the use of individual GDP or per capita GDP (also commonly employed in the literature), the bilateral product of GDP (Gdp) and difference in per capita income (SqIncome) reduces both multicollinearity with country fixed effects and problems of identification of income elasticities (Santos-Silva and Tenreyro, 2006).

[7]In the final estimated model, however, some of the exporter fixed effects had to be dropped to avoid collinearity problems, in particular with export specific variables such as 'landlocked', or import specific, such as 'internal distance'.

4. MAIN RESULTS

The sectors candidates to find a more relevant 'small share' problem are those where small initial shares are combined with high tariffs. In Table 3, a description of average tariff (ad-valorem) equivalent[8] and markets shares for the years 2001, 2004, 2007 and 2011, in the routes from Mercosur to EU, and from EU to Mercosur, are presented. In the route from Mercosur to EU, tariff equivalents higher than 40% are found in cattle meat, dairy, processed rice and sugar, while trade shares (with the exception of cattle meat, 11%) are lower than 5%(sugar) and even 1% (dairy and rice). In the route EU to Mercosur, it's more difficult to find a clear candidate, as the combination of high tariff/low share is not so straightforward.

TABLE 3 . TARIFF EQUIVALENTS AND TRADE SHARES (mean 2001-2011, 2011 in brackets)

Sector	OBSERVED IMPORT SHARES (% MSHRS)		Tariff Equivalents (%)	
	Bilateral route		Bilateral route	
	Mercosur → EU	EU → Mercosur	Mercosur → EU	EU → Mercosur
Wheat	0.3[0.3]	0.1[0.0]	7.1[1.4]	3.4[0.1]
Vegetables, fruits and nuts	3.4[3.00]	7.9[10.4]	8.2[6.5]	7.3[2.0]
Meat of cattle	10.8[8.5]	4.3[2.4]	67.5[35.4]	7.6[0.9]
Other meat products	5.0[4.6]	11.7[10.5]	21.7[13.7]	9.7[2.7]
Vegetable oils and fats	22.3[24.4]	26.2[23.9]	2.9[0.1]	8.8[3.6]
Dairy products	0.0[0.0]	16.7[9.4]	41.4[2.3]	13.7[5.5]
Processed rice	0.9[1.1]	1.1[2.2]	54.4[5.4]	8.6[0.5]
Sugar	4.9[11.7]	3.9[0.8]	79.9[26.7]	15.5[2.2]
Other food products	2.8[1.7]	15.7[15.3]	11.3[10.2]	12.6[9.1]
Beverages and tobacco products	0.7[0.8]	38.9[28.1]	11.3[5.1]	17.6[13.3]

Source: Own calculations based on GTAP, several releases.

Following Philippidis et al. (2014) and Kuiper and van Tongeren (2006), the gravity equations have been applied to simulate import shares after full removal of trade barriers,

[8] The `tariffs' taken from GTAP databases are actually tariff equivalents calculated for the different types of protection (i.e. specific, ad-valorem, mixed, compound tariffs; tariff rate quotas)

for each of the 10 agro-food sectors under study. In a first step, the gravity equation is estimated. Taking into account the parsimony and the ability to handle with over-dispersed data, Poisson family models, and in particular, Poisson and Negative Binomial have been used.

Table A.2 in the annex[9] shows that in all Poisson estimated models the (squared) correlation between observed and predicted trade is above 0.60, indicating that the fit of the model is acceptable, whilst in half of the eight Poisson models, the fit is even larger than to 0.75, which can be considerable as very good. Focusing on tariff barriers (Mt), all coefficients are negative, while in more than two thirds of the regressions, ad-valorem tariffs have a significant coefficient. These estimations serve as an input for calculating the predicted trade shares after liberalisation, as explained in the previous section.

The main simulation results after full tariff elimination between Mercosur and the EU (bi-directional trade) are presented in Table 4. An inspection of trade data reveals that in 'wheat' (wht) and 'cattle meat' (cmt) sectors, tariffs imposed by Mercosur to the EU are almost null in 2011, and therefore, the liberalisation process appears to have advanced substantially in the period 2001/2011. The same process may be occurring in 'vegetable oils and fats' (vol) sector, but in this case in the opposite direction, from Mercosur to the EU. Consequently, no relevant predicted import percentage changes were obtained in these cases (from EU to Mercosur: 0.7 and 1.2% in 'wht' and 'cmt' sectors, respectively; from Mercosur to EU: 0.2% in 'vol' sector).

Distinguishing between bilateral routes, when the destination region is the EU, the import percentage change after tariff elimination ranges from 0.2 to 53.9% in 'vol' and 'mil' sectors, respectively (when the destination region is Mercosur, from 0.7 to 24.3% in 'wht' and 'ofd' (other food products) sectors, respectively).

After eliminating tariff barriers, remarkable changes in Mercosur to EU trade flows have been found in those sectors where high tariffs were originally imposed (in year 2011), as

[9] A common tariff coefficient has been estimated for all bilateral trade routes under study. Specific tariff coefficients for each bilateral trade route has also been calculated, however, this more flexible specification was finally discarded as counterintuitive positive tariff coefficients were obtained in some regressions.

occurs with red meat(cmt, 35%), white meat (omt, 14%), and ‘sugar’ (sgr 27%). In those sectors, mean market shares in the period 2001-2011 were: 11% in red meat, and 5% in white meat and sugar, and the gravity prediction leads to substantial changes in market changes after tariff removal: 7.5 (cmt), 3.00 (omt), and 35.2(sgr), respectively (see table 4). Other sectors that experiment sharp percent changes in market shares from Mercosur to the EU are dairy (53.9%), other food products (30.3%) and processed rice (12.1%). However, in these sectors, original market shares were very small (under 1% in dairy and processed rice) and around 3% in other food products, so final market shares will still remain quite small.

TABLE 4OBSERVED MEAN (2001-2011) IMPORT SHARES AND THEIR PERCENTAGE CHANGE AFTER FULL TARIFF ELIMINATION

Sector	**OBSERVED IMPORT SHARES (% MSHRS)**		**% CHANGE AFTER TARIFF REMOVAL (% GSHRS/MSHRS)**	
	Bilateral route		**Bilateral route**	
	Mercosur → EU	EU → Mercosur	Mercosur → EU	EU → Mercosur
Wheat	0.3	0.1	17.0	0.7
Vegetables, fruits and nuts	3.4	7.9	3.7	3.4
Meat of cattle	10.8	4.3	7.5	1.2
Other meat products	5.0	11.7	3.0	1.2
Vegetable oils and fats	22.3	26.2	0.2	7.0
Dairy products	0.04	16.7	53.9	14.0
Processed rice	0.9	1.1	12.1	3.5
Sugar	4.9	3.9	35.2	8.0
Other food products	2.8	15.7	30.3	24.3
Beverages and tobacco products	0.7	38.9	7.8	5.0

5. FINAL REMARKS

The Armington specification of trade demand functions embedded in many General and Partial Equilibrium models understates trade creation effects resulting from tariff reductions in those (potentially competitive) countries whose exports typically have small trade shares in the import portfolio. The report illustrates how to use the gravity equation to predict trade shares after tariff reduction or elimination in a particular bilateral route. These new trade shares can then be used in combination with original data to shift the Armington (1969) trade demand function in the predicted direction, in order to mitigate the ´small shares stay small´ problem present not only in CGE but also other economic models which use the Armignton function.

The econometric estimation of trade shares following tariff reduction or elimination can be adapted to the specific needs of the modeller aiming at a better consistency of the data between the econometric estimation and the partial/general equilibrium model used. Thus, the method is flexible enough to accommodate alternative geographical and sectorial aggregations, as well time periods, as far as data are available. Nevertheless, the database used in the estimation needs to be large enough to get good estimates, in particular for the tariff coefficient, and guarantee sufficient degrees of freedom.

In this report, the example of bilateral trade liberalization between Mercosur and the EU is considered, in ten agro-food sectors.

Some of the sectors where Mercosur originally enjoys small market shares among EU imports, as the ones considered in this report (with the exception of vegetable oils), experiment substantial changes in their market shares (in percent terms). Note that trade is only affected on the route where the tariff is shocked. The gravity model therefore only provides a first order approximation of trade creation changes to the beneficiary country (i.e., Mercosur). These are subsequently inputted into the CGE model as import demand shifters for Mercosur trade (based on predicted gravity trade shares) whilst the CGE model takes care of the resulting (additional) impacts on trade diversion for other sources (eg. the EU and Rest of the World).

REFERENCES

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

Anderson, J.E., van Wincoop, E. (2004). Trade Costs. *Journal of Economic Literature* 42, 691-751.

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

Arnon A., Spivak A., Weinblatt J. (1996). The potential for trade between Israel, the Palestinians and Jordan. *The World Economy* 19(1), 113-134.

Baier S.L., Bergstrand, J.H. (2009). Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation. Journal of International Economics 77, 77-85.

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

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

Cameron, A.C., Trivedi, P.K. (2010). Microeconometrics using STATA. Revised Edition. STATA Press, StataCorp LP, College Station, Texas, US.

CEPII - Centre d'Études Prospectives et d'Informations Internationales (2014). (available at: www.cepii.fr).

Dimaranan, B.V. (Ed.) (2006). Global Trade, Assistance, and Production: The GTAP 6 Data Base, Center for Global Trade Analysis, PurdueUniversity, West Lafayette, IN.

Engelbert, T., Bektasoglu, B., Brockmeier, M. (2014). Moving toward the EU or the Middle East? An assessment of alternative Turkish foreign policies utilizing the GTAP framework. *Food Policy* 47, 46-61.

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

Foster N., Poeschl J., Stehrer R.(2011). The impact of Preferential Trade Agreements on the margins of international trade. *Economic Systems* 35, 84-97.

Gourieroux C., Monfort, A., Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson models. *Econometrica* 52, 701-720.

Hallack J.C. (2006). Product quality and the direction of trade. *Journal of International Economics* 68, 238-265.

Hayakawa K., Yamashita N. (2011). The role of preferential trade agreements (PTAs) in facilitating global production networks. *Journal of World Trade* 45, 1181-1207.

Kandogan Y. (2008). Consistent estimates of regional blocs' trade effects. *Review of International Economics* 16, 301-314.

Kuiper M., 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).

Linder S. (1961). An essay on trade and transformation. AlmqvistWiksell, Stockholm.

Mayer T. and Zignago S. (2011). Notes on CEPII's distances measures: The GeoDist database. CEPII Working Paper No 2011-25.

MacPhee, C., Cook, P., Sattayanuwat, W. (2013). Transportation and the international trade of Eastern and Southern Africa. *South African Journal of* Economics 81(2), 225-239.

Melitz J. (2007). North, South and distance in the gravity model. *European Economic Review* 51, 971-991.

Narayanan B. G., L. Walmsley T.L. (Eds) (2008). *Global Trade, Assistance, and Production: The GTAP 7 Data Base*, Center for Global Trade Analysis, Purdue University.

Narayanan, B. G., Badri, A. A.,McDougall, R. (Eds) (2012). Global Trade, Assistance, and Production: The GTAP 8 Data Base, Center for Global Trade Analysis, Purdue University

Philippidis G., Resano-Ezcaray H., Sanjuán-López A.I. (2014). Shifting Armington Trade Preferences: A re-examination of the Mercosur-EU negotiations. *Economic Modelling* 40, 21-32.

Pullianen K.A. (1963). World trade study: An econometric model of the patterns of the commodity flows in international trade 1948-1960. *EkonSambundetsTidskrift* 2, 78-91.

Rose A. K., van Wincoop E. (2001). National Money as a Barrier to International Trade: The Real Case for Currency Union. *American Economic Review Papers and Proceedings*, 91, 386-390.

Samuelson P. A. (1952). The transfer problem and transport costs: The terms of trade when impediments are absent. *Economic Journal*, **62**, 246, 278–304.

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

Santos Silva J.M.C., Tenreyro S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economic Letters* 112, 220-222.

Thoumi F. (1989). Trade flows and economic integration among the LDCs of the Caribbean Basin. *Social and Economic Studies* 38, 215-233.

Tinbergen J. (1962). Shaping the world economy: Suggestions for and international economic policy. Twentieth Century Fund, New York.

Woolridge J.M. (2002). Econometric Analysis of Cross Section and Panel Data. Boston, MIT Press.

World Bank (2014). World Development Indicators (http://data.worldbank.org/data-catalog/world-development-indicators 2014).

APPENDIX

TABLE A.1 LIST OF COUNTRIES USED IN ESTIMATION

Albania	Egypt, Arab Rep.	Malta	South Africa
Argentina	Estonia	Mauritius	Spain
Australia	Finland	Mexico	Sri Lanka
Austria	France	Morocco	Sweden
Bangladesh	Germany	Mozambique	Switzerland
Belgium	Greece	Netherlands	Taiwan
Bolivia	Hungary	New Zealand	Tanzania
Botswana	India	Nigeria	Thailand
Brazil	Indonesia	Pakistan	Tunisia
Bulgaria	Iran, Islamic Rep.	Paraguay	Turkey
Cambodia	Ireland	Peru	Uganda
Canada	Italy	Philippines	United Kingdom
Chile	Japan	Poland	United States
China	Korea, Dem. Rep.	Portugal	Uruguay
Colombia	Latvia	Romania	Venezuela, RB
Croatia	Lithuania	Russian Federation	Vietnam
Cyprus	Luxembourg	Senegal	Zambia
Czech Republic	Madagascar	Singapore	Zimbabwe
Denmark	Malawi	Slovak Republic	
Ecuador	Malaysia	Slovenia	
+ 16 geographical aggregations			

TABLE A.2. ESTIMATED PARAMETERS OF THE GRAVITY EQUATION WITH THE POISSON FAMILY MODELS [a, b]

Sector code [c]		Mt [d]	Dist	DistInt	NoSo	Remote	Lock	Contig	Lang	Col	PTA	Sqincome	Constant	Pseudo-R^2 [e]
wht [f]	**coef**	-0.736**	-1.694***	-4.826*	0.036	-0.100***	-1.537	-0.046	0.577***	0.244	1.185***	0.006	19.120***	0.70
	std.err	0.309	0.107	2.927	0.054	0.024	1.082	0.184	0.201	0.158	0.256	0.030	16.190	
v_f	**coef**	-0.386***	-1.315***	-2.978***	0.203***	0.026***	-0.247***	0.251***	0.381***	0.234***	0.319***	0.034	0.977	0.80
	std.err	0.024	0.004	0.518	0.002	0.001	0.070	0.006	0.008	0.006	0.009	0.000	2.891	
cmt [f]	**coef**	-0.141	-1.049***	-3.113	0.090	-0.030	1.560***	-0.020	0.436*	0.403**	1.195***	-0.022***	0.977	0.67
	std.err	0.396	0.143	2.054	0.072	0.028	0.772	0.153	0.248	0.186	0.263	0.023	11.410	
omt [b]	**coef**	-0.122	-1.138***	-1.637	0.020	0.077***	0.680	1.413***	0.396***	0.408***	0.634***	0.026***	-5.279	-
	std.err	0.181	0.051	1.722	0.030	0.009	0.511	0.147	0.152	0.122	0.095	0.011	9.581	
vol [b]	**coef**	-1.878***	-1.635***	0.879	-0.086***	0.028	0.622*	1.240***	0.454***	0.771***	-0.252***	-0.009	-13.826	-
	std.err	0.313	0.058	1.718	0.033	0.020	0.411	0.150	0.166	0.116	0.116	0.011	9.580	
mil	**coef**	-0.936***	-1.051***	-3.356	0.035	0.011*	-0.037	0.560***	0.398***	0.168	0.708***	-0.006	4.754	0.77
	std.err	0.260	0.096	2.403	0.048	0.007	0.529	0.184	0.152	0.164	0.148	0.013	13.539	
pcr	**coef**	-0.301***	-1.372***	-4.760***	-0.152***	0.151***	-1.982***	-0.232	0.335	-0.268	0.996***	-0.062***	12.594	0.76
	std.err	0.142	0.172	2.146	0.074	0.024	0.960	0.209	0.332	0.248	0.292	0.029	12.154	
sgr	**coef**	-0.520***	-1.363***	-8.826***	0.202***	0.042***	2.064***	0.595***	0.606***	0.448***	0.585***	-0.020	35.898***	0.68
	std.err	0.237	0.122	2.132	0.077	0.019	0.506	0.191	0.251	0.159	0.256	0.021	12.037	
ofd	**coef**	-2.224***	-0.869***	-4.019*	-0.015	0.021***	-0.290	0.575***	0.226***	0.216***	0.171*	0.032***	9.924	0.84
	std.err	0.388	0.048	2.323	0.033	0.004	0.392	0.086	0.112	0.102	0.093	0.011	12.993	
b_t	**coef**	-0.479**	-0.890***	-1.046	0.115***	0.026***	0.373	0.214	0.593***	0.256*	0.414***	-0.014	-8.485	0.71
	std.err	0.294	0.083	2.155	0.058	0.008	0.464	0.204	0.164	0.142	0.176	0.017	12.196	

Notes to Table A.1:

[a] Results for the year and country-specific fixed effects are not reported for space saving reasons; ***, ** and * stand for significant coefficients at 1, 5 and 10% of level of significance, respectively.

[b] Despite over-dispersion has been detected in all sectors, the Poisson model, with robust standard errors, has been applied with the exception of the 'omt' and 'vol' sectors, where the NB was used as provided more sensible tariff coefficients. Moreover, over-dispersion parameter (α)was found statistically different from zero value in both 'omt' (Chi squared statistic: 161908.12, p-value: 0.000)and 'vol' sectors (Chi squared statistic: 330696.24, p-value: 0.000)following Likelihood ratio test between NB and Poisson nested models (Ho: α = 0), and highly significant ('omt'- over-dispersion coefficient: 1.527, t-test: 31.412, p-value: 0.000; 'vol'- over-dispersion coefficient: 2.480, t-test: 39.129, p-value: 0.000). See footnote 5 in the report.

[c] The sector codes are described in Table 1.

[d] The variable abbreviations are described in Table 2.

[e] Square correlation between actual and fitted values is normally used in count models as a fit measure given the lack of a closer equivalent to the R2 in linear models (Cameron and Trivedi, 2010).

[f] 2011 year data excluded from estimation yielding 26838 observations instead of 35784. A large amount of zero tariffs (wht: 89.4%, cmt: 80.7% of observations in 2011) influenced the tariff sign, rendering it both negative and extremely small (wht) or positive (cmt).

European Commission
EUR 27264 – Joint Research Centre – Institute for Prospective Technological Studies

Title: **Gravity estimations to correct the ´small shares stay small´ bias in economic models.**

Author(s): Ana Isabel Sanjuán, Helena Resano
Luxembourg: Publications Office of the European Union

2015 – 23 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)
ISBN 978-92-79-48391-2 (PDF)
doi:10.2791/531517

LF-NA-27264-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

doi:10.2791/531517

ISBN 978-92-79-48391-2

Publications Office