# Comparing alternative methods to estimate gravity models of bilateral trade

**Estrella Gómez Herrera.** Department of Economic Theory, University of Granada, Granada, Spain (e-mail: estrellagh@ugr.es)

**Abstract**. The gravity equation has been traditionally used to predict trade flows across countries. However, several problems related with its empirical application still remain unsolved. The unobserved heterogeneity, the presence of heteroskedasticity in trade data or the existence of zero flows, which make the estimation of the logarithm unfeasible, are some of them. This paper provides a survey of the most recent literature concerning the specification and estimation methods of this equation. For a dataset covering 80% of world trade, the most widely extended estimators are compared, showing that the Heckman sample selection model performs better overall for the specification of gravity equation selected.

***Keywords*** *International trade, gravity model, estimation methods*

JEL Classification C13, C33, F10

## 1. Introduction

The gravity model of trade, which was originally inspired by Newton's gravity equation, is based on the idea that trade volumes between two countries depend on their sizes in relation to the distance between them. In the last fifty years, this model has been widely used to predict trade flows.

The gravity equation appears to be highly effective at this point as proven at a very early date by the works of Linnemann (1966) and Leamer and Stern (1971). However, several controversies have arisen regarding the model. The theoretical framework was put into doubt and afterwards justified by Bergstrand (1989) for the factorial model, Deardorff (1998) for the Hecksher-Ohlin model, Anderson (1979) for goods differentiated according to their origin, and Helpman et al. (2008) in the context of firm heterogeneity. After some additional discussions concerning its specification in the nineties, the debate has now turned to the performance of different estimation techniques. New estimation problems concerning the validity of the log linearisation process of the gravity equation in the presence of heteroskedasticity and the loss of information due to the existence of zero trade flows have been recently explored.

Traditionally, the multiplicative gravity model has been linearised and estimated using OLS assuming that the variance of the error is constant across observations (homoskedasticity), or using panel techniques assuming that the error is constant across countries or country-pairs. However, as pointed out by Santos Silva and Tenreyro (2006), in the presence of heteroskedasticity, OLS estimation may not be consistent and nonlinear estimators should be used. Another challenge described in the literature concerns the zero values. Helpman et al. (2008) propose a theoretical foundation based on a model with heterogeneity of firms à la Melitz (2003) and an adapted Heckman procedure to predict trade taking into account these features. Recently, the works of Burger et al. (2009), Martin and Pham (2008), Martínez-Zarzoso et al. (2007), Siliverstovs and Schumacher (2009) and Westerlund and Wilhelmsson (2009) have obtained divergent results when comparing alternative estimation methods.

This paper reviews most estimation methods and problems and provides a survey of the literature related to this topic. The performance of several linear and nonlinear estimators is compared using a three-dimensional (*i, j, t*) dataset, analysing the most relevant properties of each one. To this end, a gravity equation based on Anderson and van Wincoop's (2003) theoretical model is used. Using this equation, the fit of different estimation procedures applied to a large dataset of bilateral exports for 80 countries (80% of world trade) over the 1980-2008 period is discussed. The fit of each method is compared through different measures, revealing the main advantages and disadvantages of each one. It is shown that methods that do not properly treat the presence of zero flows on data exhibit noticeably worse performance than others. On the other hand, nonlinear estimators show more accurate results and are robust to the presence of heteroskedasticity in data. Overall, the Heckman sample selection model is revealed to be the estimator with the most desirable properties, confirming the existence of sample selection bias and the need to take into account the first step (probability of exporting) to avoid the inconsistent estimation of gravity parameters.

The rest of the paper is organised as follows. The next section briefly reviews the different theoretical foundations of the gravity equation to justify the election of the empirical specification of the gravity equation chosen. Section 3 compares in detail the different estimation methods available in the gravity literature. In Section 4, data are presented and the results of different estimations methods are discussed and compared to different criteria. Conclusions are drawn in Section 5. The figures and tables are provided in the Appendix.

## 2. The gravity equation

The theoretical foundation of the gravity equation appeared seventeen years after its empirical specification. The first article providing a microfoundation of this equation was Anderson (1979) and was based on the Armington assumption of specialisation of each nation in the production of only one good. Bergstrand (1985) initially supported this hypothesis, completing the theoretical foundation with a more detailed explanation of the supply side of economies and the inclusion of prices in the equation.

A few years later, a new wave of developments came with what has been called "the new trade theory". The main improvement was the replacement of the assumption of product differentiation by country of origin by the assumption of product differentiation among producing firms. In this line, Bergstrand (1990) provided a foundation based on Dixit and Stiglitz's monopolistic competition assumption. In addition, he generalised the model by introducing prices and incorporating the Linder hypothesis. Helpman (1987) also derived a foundation relying on the assumption of increasing returns to scale where products were differentiated by firms, not only by country, and firms were monopolistically competitive. However, some years later Deardoff (1998) asserted that the gravity equation could be derived from standard trade theories, conciliating both the old and the new theories.

Later on, the "new new trade theory" insisted on the heterogeneity of firms regarding their exporting behaviour (Melitz 2003), thereby giving a theoretical foundation for the presence of zero trade flows in data. In this line, Helpman et al. (2008) generalised the empirical gravity equation by developing a two-stage estimation procedure that takes into account extensive and intensive margins of trade. They showed that the incorrect treatment of zero flows may lead to biased estimates and developed a complete framework to provide a rationale for the existence of these flows.

Regarding the specification, Anderson and van Wincoop (2003) propose an augmented version of the Anderson (1979) model based on the assumption of differentiation of goods according to place of origin. Their main contribution is the inclusion of multilateral resistance terms for the importer and the exporter that proxy for the existence of unobserved trade barriers. This model is interesting overall to the extent that the discussion of the multilateral resistance may matter for heteroskedasticity considerations. In this model, countries are representative agents that export and import goods. Goods are differentiated by place of origin and each country is specialised in the production of only one good. Preferences are identical, homothetic and approximated by a constant elasticity of substitution (CES) function.

The linear gravity equation estimated by Anderson and van Wincoop is as follows:

$$\ln X_{ij} = k + \ln y_i + \ln y_j + (1-\sigma)\rho \ln d_{ij} + (1-\sigma)\ln b_{ij} + (1-\sigma)\ln P_i + (1-\sigma)\ln P_j + \varepsilon_{ij} \qquad (1)$$

where $X_{ij}$ is the nominal value of exports from $i$ to $j$; $k$ is a positive constant, $y_i$ and $y_j$ are the nominal income of each country, generally proxied by its GDP, and $d_{ij}$ is a measure of the bilateral distance between $i$ and $j$, which are introduced to proxy for transport costs. $b_{ij}$ is a dummy variable that takes value one if two countries share a border. Finally, the variables $P_i$ and $P_j$ are the multilateral resistance terms and are defined as a function of each country's full set of bilateral trade resistance terms. The variable of interest for Anderson and van Wincoop is $b_{ij}$ since their objective is to estimate the trade effect of national borders. They apply their equation to regional data.

The multilateral price indices ($P_i$ and $P_j$) are not observed and should be estimated. Anderson and van Wincoop (2003) use the observed variables in their model (distances, borders, and income shares) to obtain the multilateral trade resistance terms. Assuming symmetric trade costs, using 41 goods market-equilibrium conditions[1] and a trade cost function defined in terms of observables, they obtain the $P_i$ and $P_j$ terms. Although they argue that this method is more efficient than any other, it is highly data consuming and has not been frequently used by other authors.

An alternative solution is to include a remoteness variable to proxy for these multilateral trade resistance indexes:

$$Rm_i = \sum_j \frac{d_{ij}}{(y_j / y_{ROW})} \qquad (2)$$

where the numerator would be the bilateral distance between two countries, and the denominator would be the share of each country's GDP in the rest of the world's GDP. Head and Mayer's (2000) remoteness variable describes the full range of potential suppliers to a given importer, taking into account their size, distance and relevant costs of crossing the border. Wei (1996), Wolf (1997), and Helliwell (1996) provide other examples of regressions including a remoteness variable. Alternatively, Feenstra (2002) proposes introducing importer and exporter fixed effects to account for the specific country multilateral resistance term. The coefficient of the dummies for the importer

[1] Their sample contains the same 30 US states and 10 Canadian provinces that McCallum (1995) includes. There are 20 additional states, plus Columbia, which they aggregate into one. Hence, they finally have 41 equations.

and the exporter should reflect the multilateral resistance for each country. Several studies using this approach are described in the Appendix (Table A1). Finally, Baier and Bergstrand (2009) suggest generating a linear approximation of the $P_i$ and $P_j$ terms by means of a first-order Taylor series expansion.

Concerning the proxy for supply and demand sizes, the common practice is to use importer's and exporter's GDP correspondingly. In some cases GDP per capita is also introduced as a proxy for capital-labour intensities.

Transaction costs are frequently proxied by geographical distance. However, it is commonly accepted that geographical distance may be a poor approximation[2]. Thus, this variable is often completed with other proxies for trade barriers specified as indicator variables. For instance, adjacency takes value one if trade partners share a common border, common language takes value one if both countries share a language, colonial links captures the effect of having had a common coloniser or having been colonised by another country in the past; religion takes value one when both countries have the same religion; access to water takes value one if a country has access to water, or Regional Trade Agreement (RTA) which assess the effect of RTAs on trade. All these factors affect international trade via transaction costs and complete the geographical distance variable in order to reflect the economic distance.

# 3. Summary of estimation methods

As mentioned above, interest in the last years has focused on estimation methods to accurately predict trade flows. In this section, a brief summary of some of the most important estimation methods as well as a revision of related empirical literature (Table 1) are presented.

## 3.1. Linear methods

Since the logarithm of zero is not defined, truncation and censoring methods have been proposed in the literature to treat the problem of zero flows in data. However, these procedures reduce efficiency due to the loss of information and may lead to biased estimates due to the omission of data. Furthermore, as Westerlund and Wilhelmsson (2009) point out, the elimination of trade flows when zeros are not randomly distributed leads to sample selection bias.

In addition, a panel framework permits recognising how the relevant variables evolve through time and identifying the specific time or country effects. Over the last years, researchers such as Egger (2000), Rose and van Wincoop (2001), Mátyás (1998), Egger and Pfaffermayr (2003, 2004), Glick and Rose (2002), Brun et al. (2002), and Melitz (2007) have turned towards panel data[3]. Two main techniques are employed to fit data depending on the a priori assumptions. The fixed effects estimator assumes the existence of an unobserved heterogeneous component that is constant over time and which affects each individual (pair of countries) of the panel in a different way. By contrast, the random effects model imposes no correlation between the individual effects and the regressors, implicitly assuming that the unobserved heterogeneous

[2] In addition, there is no single opinion about how distance should be measured. The most common measures are the great circle formula and the distance between the two principal cities. See Wei (1996), Wolf (1997), and Head and Mayer (2000) for further information.
[3] See Appendix A for further information.

component is strictly exogenous. Under the null hypothesis of zero correlation, the random effects model is more efficient. However, if the null is rejected, only the fixed effects model provides consistent estimators[4].

### 3. 2. Nonlinear methods

As Santos Silva and Tenreyro (2006) points out, the log-linearisation of the gravity equation changes the property of the error term, thus leading to inefficient estimations in the presence of heteroskedasticity. If the data are homoskedastic, the variance and the expected value of the error term are constant but if they are not -as usually happens with trade data-, the expected value of the error term is a function of the regressors. The conditional distribution of the dependent variable is then altered and OLS estimation is inconsistent. Heteroskedasticity does not affect the parameter estimates; the coefficients should still be unbiased, but it biases the variance of the estimated parameters and, consequently, the t-values cannot be trusted. Hence, the recent literature concerning estimation techniques have opted to use nonlinear methods as well as two parts models for estimating the gravity equation.

Among nonlinear estimation methods, the most frequently used are Nonlinear Least Squares (NLS), Feasible Generalised Least Squares (FGLS), Heckman sample selection model and Gamma and Poisson Pseudo Maximum Likelihood (GPML and PPML). Santos Silva and Tenreyro (2006) claim that NLS is inefficient since it gives more weight to observations with larger variance and is not robust to heteroskedasticity. Martínez-Zarzoso et al. (2007) propose Feasible Generalised Least Squares (FGLS) as the most appropriate model if the exact form of heteroskedasticity in data is ignored since it weighs the observations according to the square root of their variances and is robust to any form of heteroskedasticity. Manning and Mullahy (2001) propose Gamma Pseudo Maximum Likelihood (GPML). In this case the conditional variance of the dependent variable is assumed to be proportional to its conditional mean. This estimator therefore assigns less weight to observations with a larger conditional mean. Martínez-Zarzoso et al. (2007) computes the performance of this estimator, finding it to be adequate in the presence of heteroskedasticity, although it shows less accuracy when zero trade values are present. Finally, Poisson Pseudo Maximum Likelihood (PPML) is similar to GPML, but assigns the same weight to all observations. Santos Silva and Tenreyro (2006) point out that this is the most natural procedure without any further information on the pattern of heteroskedasticity.

In addition, two-step estimation methods have also been proposed to estimate the gravity equation. This is the case of Heckman sample selection model. In the first step, a Probit equation is estimated to define whether two countries trade or not and in a second step, the expected values of the trade flows, conditional on that country trading, are estimated using OLS. In order to identify the parameters on both equations, a selection variable is required. This exclusion variable should affect only the decision process; hence, it should be correlated with a country's propensity to export but not with its current levels of exports. Some examples in the literature are the common language and common religion variable (Helpman et al. 2008), governance indicators of regulatory quality (Shepotylo 2009), or the historical frequency of positive trade

[4] The Hausman test provides a method for testing the adequacy of the random effect model. If the null is rejected, the random effects model is not consistent. However, it is important to note that this result does not imply that the fixed effect model is adequate.

between two countries (Bouet et al. 2008). Alternatively, Linders and de Groot (2006) or Haq et al. (2010) include the same variables in both equations, imposing the normality of the error in both equations as an identification condition, which implies a zero covariance between them. The advantage of a sample selection model comes from the fact that the decision on whether to trade or not and the decision on how much to trade are not modelled as completely independent. The model allows for some positive correlation between both error terms to better reflect the real decision process. For further information on this topic see Egger et al. (2011).

Helpman et al. (2008) extends Heckman's estimation method to also take into account the bias associated with the heterogeneity of firms. The authors develop a complete theoretical framework from which they obtain an empirical specification of the gravity equation. Their model accounts for firm heterogeneity, trade asymmetries and fixed trade costs, suggesting that the decision to export (extensive margin) and the volume of exports (intensive margin) are not independent variables. The model allows both positive and zero trade flows between countries to be predicted and it also allows exports to vary according to the destination country. Helpman et al. (2008) describe a varying distribution of firms where each firm is bounded by a marginal exporter who breaks even by exporting to another country. The underlying idea is that if at least one firm in the country is productive enough to export, country-level exports in that case will be positive. Hence, zero exports are originated by countries where firms are not productive enough to export profitably. In this manner, information that would normally require firm-level data is extracted from country-level data.

They argue that controlling for both the extensive margin and the sample selection would completely eliminate the bias in the estimation. The results confirm their theoretical predictions, showing that the omission of a measure of firms' heterogeneity leads to substantial biases in the estimation. They prove the robustness of their results using religion instead of common language as exclusion variable. Most articles employing the Helpman et al. (2008) methodology apply it to a cross-section dataset. Application of the methodology in a panel framework still requires further research and goes beyond the scope of this article.

Every method has advantages and disadvantages and it cannot be asserted that any one of them absolutely outperforms the others. For that reason, it has become a frequent practice in the literature to include several estimation methods for the same database. In the next section, an empirical exercise comparing these methods is presented.

**Table 1**- Summary of estimation methods

| Estimation method | Advantages | Disadvantages | References |
|---|---|---|---|
| Truncated OLS | - Simple | - Loss of information (elimination of zero flows)<br>- Biased coefficients | Linders and de Groot (2006); Westerlund and Wilhelmsson (2009); Martin and Pham (2008) |
| OLS ($1+T_{ij}$) | - Simple<br>- It deals with the zero trade flows problem | - Biased coefficients | Linneman (1966), Bergeijk and Oldersma (1990);<br>Wang and Winters (1991); Baldwin and DiNino (2006) |
| Tobit (censored regression) | - Simple<br>- It deals with the zero trade flows problem | - Same set of variables to determine the probability that an observation will be censored and the value of the dependent variable<br>- Lack of theoretical foundation | Soloaga and Winters (2001);<br>Anderson and Marcouiller (2002);<br>Baldwin and DiNino (2006);<br>Schiavo (2007); Martin and Pham (2008) |
| Panel fixed effects | - Simple<br>- It controls for unobserved heterogeneity | - Loss of information (constant terms in the regression are dropped)<br>- Elimination of zero flows<br>- Sample selection bias | Mátyás (1998); Egger (2000); Glick and Rose (2002); Egger and Pfaffermayr (2003); Micco et al. (2003); Andrews et al. (2006); Henderson and Millimet (2008) |
| Heckman two-step | - Different set of variables and coefficients to determine the probability of censoring and the value of the dependent variable<br>- No multicollinearity problems<br>- It provides a rationale for zero trade flows | - It may be difficult to find an identification restriction<br>- Exclusion variables are required | Bikker and de Vos (1992); Linders and de Groot (2006); Martin and Pham (2008) |
| PPML (Poisson Pseudo Maximum Likelihood) | - It deals with the zero trade flows problem – It provides unbiased estimates in the presence of heteroskedasticity<br>- All observations are weighted equally<br>- The mean is always positive | - It may present limited-dependent variable bias when a significant part of the observations are censored | Westerlund and Wilhelmsson (2009); Siliverstovs and Schumacher (2009); Liu (2009);<br>Shepherd and Wilson (2009); Martínez-Zarzoso et al. (2007); Santos Silva and Tenreyro (2006); An and Puttitanun (2009) |
| NLS (Nonlinear | - It deals with the zero | - It assigns more | Santos Silva and |

| | | | |
|---|---|---|---|
| Least Squares) | trade flows problem | weight to observations with a larger variance (inefficiency).<br>- Not robust to heteroskedasticity<br>- Sample selection bias | Tenreyro (2006) |
| FGLS (Feasible Generalised Least Squares) | - It deals with the zero trade flows problem<br>- It is robust to heteroskedasticity | - The variance covariance matrix should be estimated first | Martínez-Zarzoso et al. (2007) |
| GPML (Gamma Pseudo Maximum Likelihood) | - It deals with the zero trade flows problem<br>- It is robust to heteroskedasticity | - Less weight to observations with a large conditional mean (less prone to measurement errors) | Martínez-Zarzoso et al. (2007) |
| Helpman, Melitz and Rubinstein (2008) | - It provides a rationale for zero trade flows<br>- Unbiased estimates | - Difficult to estimate<br>- Additional data is required (exclusion variables) | Helpman et al. (2008); Santos Silva and Tenreyro (2008) |

# 4. Comparing estimation methods for a baseline gravity equation

The new workhorse in the estimation of the gravity equation is still unclear. Econometric estimation presents some challenges that remain unsolved as of yet. First, the exclusion of the multilateral trade resistance terms leads to biased estimates due to the omission of variables. Anderson and van Wincoop (2003) claimed that this misspecification invalidates the estimation. Second, taking logarithms and estimating by OLS in the presence of heteroskedasticity leads to inconsistent estimates as noted by Santos Silva and Tenreyro (2006). Third, there are some aspects that may differ from one country to another but are not reflected by the regressors (i.e. regulation, political factors, technology, e-business, port efficiency, etc.). This unobserved heterogeneity should be controlled for to obtain unbiased estimates. Finally, if two countries do not trade in a given year the value of their trade would be represented by a zero in the dataset. Since the logarithm of zero is unfeasible, some information would be lost. This problem is becoming more important due to the use of disaggregated data, in which over 50% of values is zero.

## 4.1. Data and model

The sample covers bilateral exports of 80 countries over the 1980-2008 period. All the countries of the EU15, the CEE new European members, and 6 Middle East and North African (MENA) countries (Morocco, Tunisia, Egypt, Turkey, Israel and Algeria) as well as most OECD countries are included. The total number of observations should be 176,960 but is reduced to 157,080 due to missing data. Data were collected from several sources, including the CHELEM-International Trade database, the CEPII database and the World Bank.[5]

For the sake of comparison, a gravity equation based on Anderson and van Wincoop's (2003) theoretical model will be used:

$$\ln X_{ijt} = \alpha_1 \ln y_{it} + \alpha_2 \ln y_{jt} + \alpha_3 contig_{ij} + \alpha_4 comla_{ij} + \alpha_5 smctry_{ij} \quad (3)$$
$$+ \alpha_6 \ln d_{ij} + \gamma_{ij} + \gamma_{it} + \gamma_{jt} + \ln \varepsilon_{ijt}$$

The dependent variable is the logarithm of the volume of exports in current dollars from country $j$ to $i$, obtained from the CHELEM-CEPII database. $\ln y_{it}$ and $\ln y_{it}$ are the logarithms of nominal GDP in each country whose effect on trade is expected to be positive. $contig_{ij}$ (Contiguity), *comla* (Common language) and *smctry* (Same country) are dummy variables that take value one when two countries share a border, a language, or were the same country in the past, correspondingly. In all cases, the coefficient is expected to be positive. $d_{ij}$ is a variable representing the geodesic distance between $i$ and $j$ and is obtained from the CEPII database. According to Egger and Pfaffermayer (2003), country pair specific fixed effects, $\gamma_{ij}$, as well as time varying fixed effects for the importer and the exporter, $\gamma_{it}$, $\gamma_{jt}$, are included in the estimation in order to capture any importer or exporter time varying characteristics. These terms correct biases that arises from the fact that we are not estimating a cross-section but a panel (see Baldwin and Taglioni 2006). Due to the inclusion of these dummies, GDP terms are dropped from the estimation. However, as first noticed by Neyman and Scott (1948), the estimation of a Tobit and Probit models with fixed effects is inconsistent due to the incidental parameter problem. Hence, fixed effects are not included in these two models.

## 4.2 Results

Before estimating equation (3), some specification tests were conducted. First, the Likelihood Ratio (LR) and the Lagrange Multiplier (LM) tests on time and individual effects were performed. In both cases, the null hypothesis of no fixed effects is rejected. The standard F-test for the joint significance of individual and time dummies confirms this result, so it can be concluded that unobserved heterogeneity is present and OLS estimation yields biased and inconsistent estimates. A simple analysis of the residuals and the fitted values confirms the presence of heteroskedasticity in the regression (see figure 2 in Appendix C). Hence, estimation with a nonlinear method is required.

Table 2 reports the estimation outcomes resulting from the different techniques employed. The dependent variable is the logarithm of exports in all cases except for Poisson regression, in which this variable is introduced in levels.

Overall, the estimation techniques seem to affect the magnitude but not the sign of the parameters for most gravity variables. As expected, both the exporter and importer GDP increases exports regardless of the estimation method used, while the distance reduces exports. Other gravity variables are also highly significant, and proximity (either in history or in space) tends to increase exports. Belonging to a Regional Trade Agreement also increases trade, although it shows a moderate effect. The main differences among estimators are revealed in the magnitude of coefficients. Whereas the Heckman and panel methods show results that are more in line with the related literature, the incorrect

[5] The CHELEM database is previously refined using a 7-step procedure. Bilateral trade data is harmonised using reports on each of the countries involved in the transaction.

treatment of zeros is observed to lead to an overestimation of coefficients in the Tobit and OLS estimation. These differences suggest the existence of a substantial bias in the estimation of the Tobit and OLS methods. On the other hand, PPML shows the lowest coefficients; a result that is in line with Santos Silva and Tenreyro (2006) and Siliverstovs and Schumacher (2009). The goodness of fit measures also reveal the existence of significant differences among the methods compared.

Columns 2 and 3 show the results for OLS adding a constant and Tobit estimates correspondingly. In both cases, the zero flows in the dependent variable are assumed to take a value of one, which is not theoretically consistent. In fact, the visual inspection of the kernel estimates reveals that Tobit coefficients are strongly biased, whereas OLS estimators present more variance than the others.

Other alternatives in the literature that do not artificially modify the dependent variable simply propose discarding the zero flows from the estimation. These are the cases reported in the first, sixth and seventh columns. The first column shows the results for the truncated OLS estimation. Most variables have the expected sign, and are highly statistically significant, though the effect of a RTA on trade is predicted to be negative, contrary to expectations. Furthermore, as mentioned before, the OLS estimation is inconsistent due to the presence of unobserved heterogeneity. Column six shows the results for the panel estimation assuming fixed effects and column seven allows the heterogeneous component to be distributed randomly. The coefficients are also significant and show the expected sign.

The last column shows the results for the PPML estimation. In this case, the dependent variable is introduced in levels instead of logarithms. Although the sign and significance are quite similar to the other estimators, PPML notably reduces the magnitude of the coefficients as well as the standard errors. Santos Silva and Tenreyro (2006) claim that this is the preferred estimation method in the presence of heteroskedasticity.

However, none of the above methods explains the presence of zero flows. Indeed, these observations are simply dropped or censored at one. Since these procedures may lead to sample selection bias when the zeros in the sample are not random, one of the alternative solutions proposed in the literature is to use a Heckman sample selection model. While other methods treat zero flows as inexistent, Heckman considers them to be unobserved. The outcomes from the first step (Probit equation) are reported in column 4. Following Helpman et al. (2008), common language is used as an excluded variable since this variable is expected to affect the probability of exporting, but not the size of exports. Column 5 reports the results for the second step. The inverse Mills ratio is highly significant, thus confirming the existence of a sample selection bias.

Several goodness-of-fit criteria have been used in order to compare estimation methods. First, the predicted over the real value of exports in a specific year (2008) is plotted for different techniques and the dispersions of the results (Figures 3 to 9 in Appendix C) are compared. Second, the graphs of the univariate kernel density estimation are examined to gain a more accurate idea of the bias and the variance of the distribution of the predicted values in each case (Figure 1). Finally, Table 3 shows the results of three goodness-of-fit functions: the bias, the mean squared error (MSE) and the absolute error loss.

**Table 2** Results for alternative estimation methods

| | Truncated OLS | OLS (1+X) | Tobit | Probit | Heckman | Panel fixed | Panel random | PPML |
|---|---|---|---|---|---|---|---|---|
| Log of exporter GDP | | | 1.431*** | 0.0907*** | | | | |
| | | | [0.024] | [0.035] | | | | |
| Log of importer GDP | | | 1.513*** | 0.104*** | | | | |
| | | | [0.023] | [0.0342] | | | | |
| Contiguity | 0.129*** | -0.482*** | 0.0462 | -0.327 | | | 0.225*** | 0.413*** |
| | [0.030] | [0.082] | [0.402] | [0.289] | | | [0.068] | [3.53e-10] |
| Common Language | 0.929*** | 2.221*** | 2.355*** | 1.606*** | | | 1.071*** | 0.244*** |
| | [0.018] | [0.049] | [0.245] | [0.175] | | | [0.052] | [3.40e-10] |
| Same Country | 0.626*** | 0.609*** | 0.609 | -0.869** | | | 0.712*** | 0.007*** |
| | [0.048] | [0.147] | [0.599] | [0.375] | | | [0.094] | [6.30e-10] |
| Log of Distance | -1.318*** | -1.943*** | -1.866*** | -0.873*** | | | -1.330*** | -0.644*** |
| | [0.008] | [0.024] | [0.074] | [0.063] | | | [0.021] | [1.59e-10] |
| RTAboth | -0.0625*** | -0.779*** | 0.0436 | 0.757*** | 0.336*** | 0.337*** | 0.292*** | 0.441*** |
| | [0.017] | [0.046] | [0.070] | [0.130] | [0.0147] | [0.0382] | [0.014] | [3.74e-10] |
| Inverse Mills Ratio | | | | | 0.617*** | | | |
| | | | | | [0.0908] | | | |
| Constant | 14.64 | -11.50*** | 11.50*** | 2.777*** | 5.314 | 16.76*** | 14.22*** | 14.91*** |
| | [] | [2.362] | [0.676] | [0.649] | [6.147] | [2.478] | [0.787] | [1.46e-07] |
| Observations | 147,954 | 157,080 | 157,080 | 157,080 | 147,954 | 147954 | 147,954 | 157,080 |

*Note:* Figures in brackets are robust standard errors. The dependent variable is the logarithm of exports in all cases except for Poisson regression, in which this variable is introduced in levels. All specifications except Tobit and Probit include importer and exporter time varying effects.
* significant at 10%; ** significant at 5%; *** significant at 1%

The main advantage to the last function, which was suggested by Martínez-Zarzoso et al. (2007), is that over- and under-estimations are not cancelled out. It is defined as follows:

$$L(X_{ijt}, \hat{X}_{ijt}) = \left| X_{ijt} - \hat{X}_{ijt} \right| \quad (4)$$

Figure 1 plots the kernel density estimates of the distributions of the predicted values from each method, as well as the observed data. The logarithm of exports is normally distributed and slightly right skewed. A one-to-one comparison of the methods reveals that almost all the estimators are slightly left skewed and present a bias with different magnitudes. The distribution of fixed PPML notably differs from all others in kurtosis (it shows a positive and high kurtosis and hence a smaller variance), whereas the rest tend to be platykurtic (higher variance). However, it exhibits a stronger bias. Hence, although it shows a smaller variance, the prediction is very poor for low trade values, which are overestimated. The plot of individual graphs for a cross-section (figures 3 to 9 in Appendix C) and the different measures of goodness of fit in Table 3 confirm this result. On the other hand, Tobit and OLS adding a constant show a very high variance, which is related to the fact that both methods treat the zeros in the sample in an incorrect manner, thus forcing the observations to have no theoretical justification. Overall, the distribution of Heckman, truncated OLS and panel random effects seem to be closest to the real distribution.

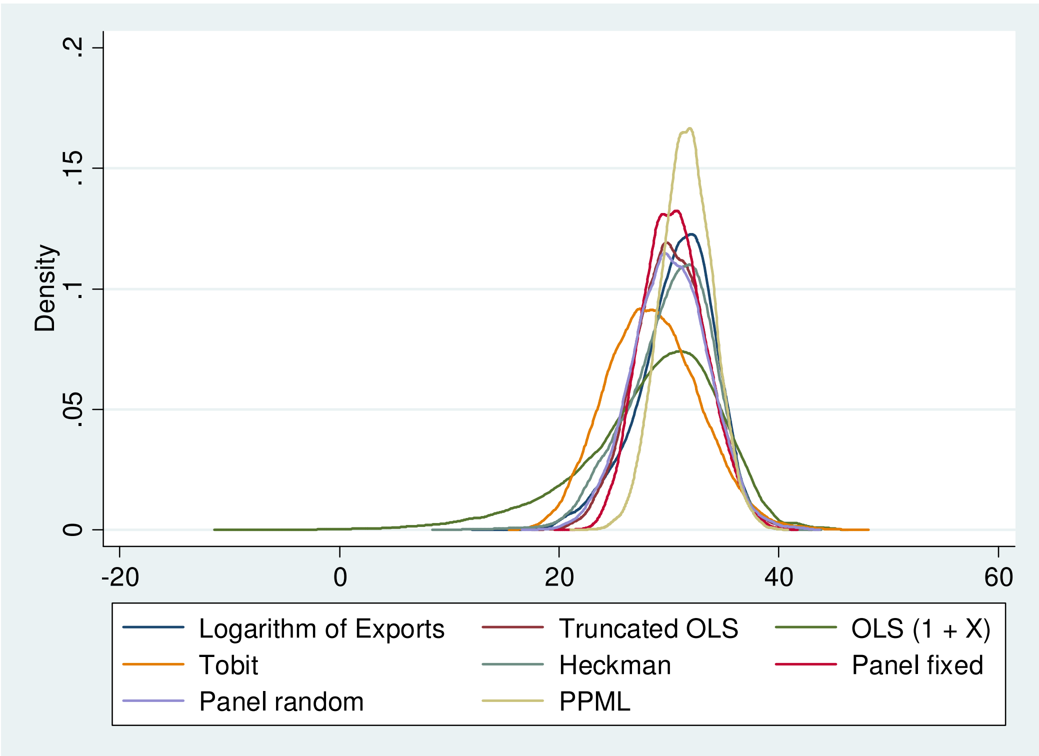


**Fig. 1** Kernel densities of different estimators

Concerning the other goodness-of-fit criteria employed, the outcomes in Table 3 confirm the abovementioned results. Heckman is the preferred estimation method regarding the MSE and absolute error loss criteria, followed by Pooled OLS and panel random effects; whereas Tobit, OLS with a modified dependent variable and panel fixed effects estimation obtain the worst results.

**Table 3** Goodness of fit

| | Bias | MSE | Error loss |
|---|---|---|---|
| Truncated | **7.95E-11** | 2.415 | 1.111 |
| OLS (1+X) | -1.069 | 9.955 | 2.200 |
| Tobit | -1.667 | 8.104 | 2.303 |
| Heckman | 8.86E-11 | **0.950** | **0.623** |
| Panel fe | -4.61E-11 | 13.315 | 2.915 |
| Panel re | -0.079 | 2.476 | 1.139 |
| PPML | 1.221 | 5.403 | 1.553 |

# 5. Concluding remarks

The gravity model is considered one of the most successful empirical frameworks in international economics. It has become a successful tool for the evaluation of trade policies or the calculation of trade potential associated with regional integration. However, a more detailed analysis of the theoretical underpinnings, the use of larger datasets and improvements in statistical and econometric software have highlighted new problems in estimating the gravity equation.

This paper has provided an in-depth review of recent developments in the literature on estimation methods for the gravity equation, finding that there are at least two problems related to the log linearisation of the gravity equation that require further research as there is no consensus about the optimal method to solve them. First, the exclusion of the multilateral trade resistance terms defined by Anderson and van Wincoop (2003), as well as the unobserved heterogeneity present in trade data leads to biased estimates due to misspecification. One usual procedure to solve this problem is to log linearise the model and to estimate it by OLS with fixed effects. However, the heteroskedasticity intrinsic to the log-linear formulation of the gravity model can result in biased and inefficient estimates when applying OLS. Second, the logarithm of zero is unfeasible. As a result, the presence of zero trade flows in data means that these observations must either be dropped or replaced by an arbitrary positive value, leading to sample selection bias and loss of information. This problem is becoming increasingly important due to the use of disaggregated datasets in which over 50% of values are zero.

An empirical exercise to compare several techniques with a dataset covering 80% of world trade has been conducted. The equation is based on the Anderson and van Wincoop (2003) specification of the gravity equation, allowing for different assumptions about the unobserved heterogeneity component. After applying several criteria to test goodness of fit, it is argued that ad hoc methods are not appropriate for estimating the gravity equation since they provide biased and inefficient estimates. On the other hand, although the use of PPML has been proposed by several authors in the literature, it does not behave so well for an aggregated dataset in the presence of unobserved heterogeneity. This paper suggests that the Heckman sample selection model is the preferred estimation method within nonlinear techniques when data are heteroskedasticity and contain a significant proportion of zero observations.

**Acknowledgments** The author is indebted to J. Milgram for her helpful comments. The author gratefully acknowledges the financial support from the Spanish *Ministerio de Ciencia e Innovación (MICINN)*, project SEJ2007-62081.

# Appendix A

Table A1 Articles using fixed effects, random effects or both effects in the estimation of the gravity equation

| **Article** | **Effects included** | **Data** | **Dependent variable** |
|---|---|---|---|
| Mátyás (1998) | - Importer, exporter and time effects | 11 countries; 1982-1994 | Exports |
| Rose and van Wincoop (2001) | - Importer, exporter and time effects | 200 countries; data at five-year intervals between 1970 and 1995 | Bilateral trade |
| Glick and Rose (2002) | - Country-pair fixed effects<br>- Symmetric country-pair effects. | 217 countries; 1948- 1997 | Real bilateral trade |
| Baltagi et al. (2003) | - Importer, exporter and time effects<br>- Country-pair fixed effects<br>- Importer-time effects<br>- Exporter-time effects | EU15, USA and Japan with their 57 most important trading partners; 1986–1997 | Real bilateral exports |
| Micco et al. (2003) | - Time effects<br>- Country-pair fixed effects | 22 developed countries; 1992 - 2002 | Bilateral trade |
| De Benedictis and Vicarelli (2005) | - Country-pair fixed effects<br>- Dynamic effects (Arellano and Bond estimator) | Each of former 11 Eurozone countries to 32 importer countries; 1991-2000 | Exports |
| Cheng and Wall (2005) | - Country-pair fixed effects<br>- Time effects | 29 countries; 1982, 1987, 1992, and 1997 | Real exports |
| Fratianni and Hoon-Oh (2007) | - Country-pair and time fixed effects<br>- Random effects | 143 countries; 1980-2003 | Real bilateral imports |
| Ruiz and Vilarrubia (2007) | - Importer, exporter and time effects<br>- Exporter-period and importer-period dummies (annual, triennial and quinquennial) | 205 countries; 1948-2005 | Bilateral trade |
| Cafiso (2008) | - Country-pair and time fixed effects | 24 OECD countries (sectors 15-37, ISIC Rev. 3); 1993-2003 | Exports |
| Fidrmuc (2008) | - Country-pair and time effects | 19 OECD countries; 1980-2002 | Bilateral trade flows |
| Henderson and Millimet (2008) | - Importer, exporter and time effects<br>- Country-pair fixed effects | US data. 25 two-digit SIC industries;1993 and 1997 | Nominal value of exports |
| Hoon-Oh and Selmier II (2008) | - Country-pair fixed effects<br>- Random effects | 859 pairs; 1980–2001 | Imports |
| Kavallari et al. (2008) | - Random effects | German imports of olive oil from 14 exporting countries; 1995-2006 | Imports |
| Bussière and Schnatz (2009) | - Country-pair fixed effects | 61 countries; 1980-2003 | Bilateral trade |
| Yu (2010) | - Fixed effects | 157 countries; 1962–1998 | Exports |

**Table A2** Articles related to the problem of zero-flows and heteroskedasticity

| Article | Data | Estimation methods | Dependent variable | Simulation studies |
|---|---|---|---|---|
| Santos Silva and Tenreyro (2006) | 136 countries; 1990 | PPML, NLS, GPML, OLS, ET-tobit, OLS ($y > 0.5$) OLS (y+1) | Trade | - PPML, NLS, GPML<br>OLS; OLS($y + 1$); truncated OLS ET-tobit.<br>- Four different patterns of heteroskedasticity |
| Martínez-Zarzoso et al. (2007) | 3 datasets:<br>1) 180 countries; 1980-2000<br>2) 47 countries; 1980-1999<br>3) 65 countries; data for every 5 years over 1980-1999 | FGLS, GPML, Poisson, Heckman | Exports | - OLS, NLS, GPML, PPML and FGLS |
| Helpman et al. (2008) | 158 countries; 1970-1997 | HMR, NLS, semi-parametric, non-parametric | Exports | No |
| Martin and Pham (2008) | 136 countries; 1990 | Truncated OLS, ET-Tobit, PPML, Heckman ML, Heckman 2SLS | Bilateral trade | - Truncated OLS, OLS (y+1), truncated NLS, censored NLS, GPML, PPML, truncated PPML, ET-Tobit, Poisson-Tobit, Heckman |
| Santos Silva and Tenreyro (2008) | 158 countries; 1986 | HMR, NLS, semi-parametric, non-parametric, GPML | Exports | No |
| Burger et al. (2009) | 138 countries; 1996-2000 | OLS, PPML, ZIPPML, BPPML | Exports | No |
| Siliverstovs and Schumacher (2009) | 22 OECD countries; 1988-1990. Disaggregated data: 25 three-digit ISIC Rev.2 industries | OLS, PPML | Trade | No |
| Westerlund and Wilhelmsson (2009) | EU and other developed countries; 1992-2002 | OLS, fixed effect PPML | Nominal imports | - OLS, truncated OLS, OLS (y+1), PPML<br>- Two patterns of heteroskedasticity |
| Yu (2010) | 157 countries 1962–1998 | OLS, fixed effects, IV, PPML | Exports | No |

# Appendix B

**Table B1** List of countries included in the sample

| | | |
|---|---|---|
| Albania | Gabon | Paraguay |
| Algeria | Germany | Peru |
| Argentina | Greece | Philippines |
| Australia | Hong Kong | Poland |
| Austria | Hungary | Portugal |
| Bangladesh | Iceland | Romania |
| Belarus | India | Russian Federation |
| Belgium and Luxembourg | Indonesia | Saudi Arabia |
| Bolivia | Ireland | Singapore |
| Bosnia and Herzegovina | Israel | Slovakia |
| Brazil | Italy | Slovenia |
| Brunei Darussalam | Japan | South Korea |
| Bulgaria | Kazakhstan | Spain |
| Cameroon | Kenya | Sri Lanka |
| Canada | Kyrgyzstan | Sweden |
| Chile | Latvia | Switzerland |
| China | Libyan Arab Jamahiriya | Taiwan |
| Colombia | Lithuania | Thailand |
| Côte d'Ivoire | Macedonia | Tunisia |
| Croatia | Malaysia | Turkey |
| Czech Republic | Mexico | Ukraine |
| Denmark | Morocco | United Kingdom |
| Ecuador | Netherlands | United States |
| Egypt | New Zealand | Uruguay |
| Estonia | Nigeria | Venezuela |
| Finland | Norway | Vietnam |
| France | Pakistan | |

# Appendix C: Cross-validation for the different estimation methods in year 2008

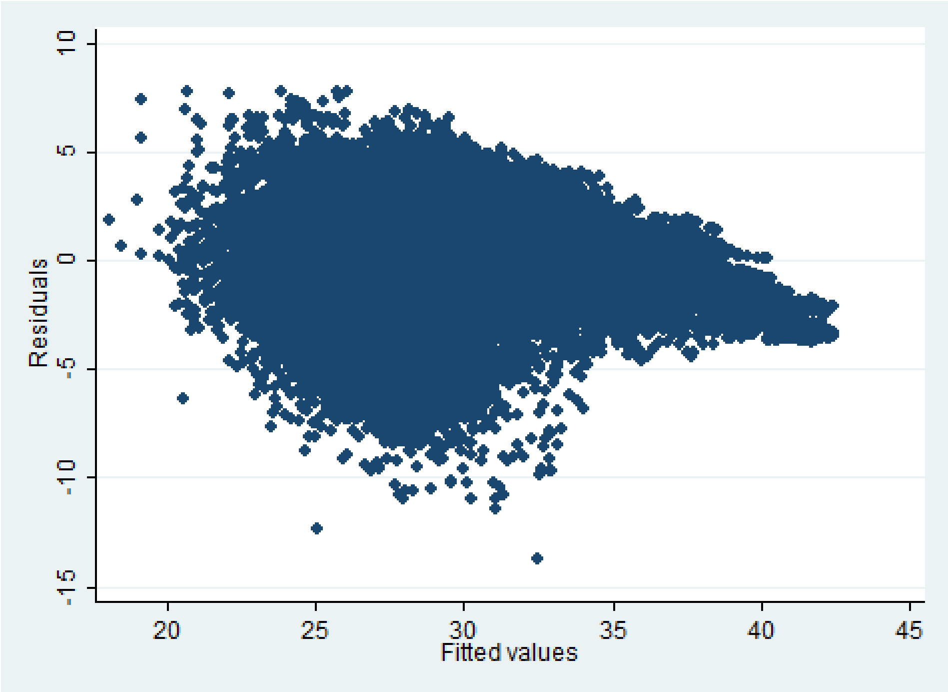


**Fig. 2** Heteroskedasticity in data. Distribution of errors

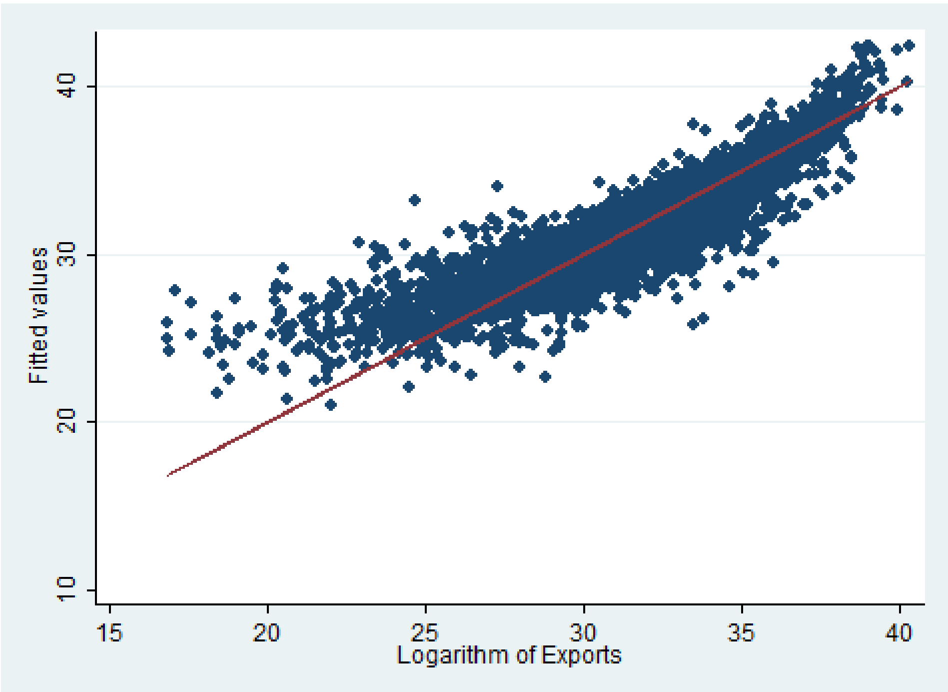


**Fig. 3** Truncated OLS

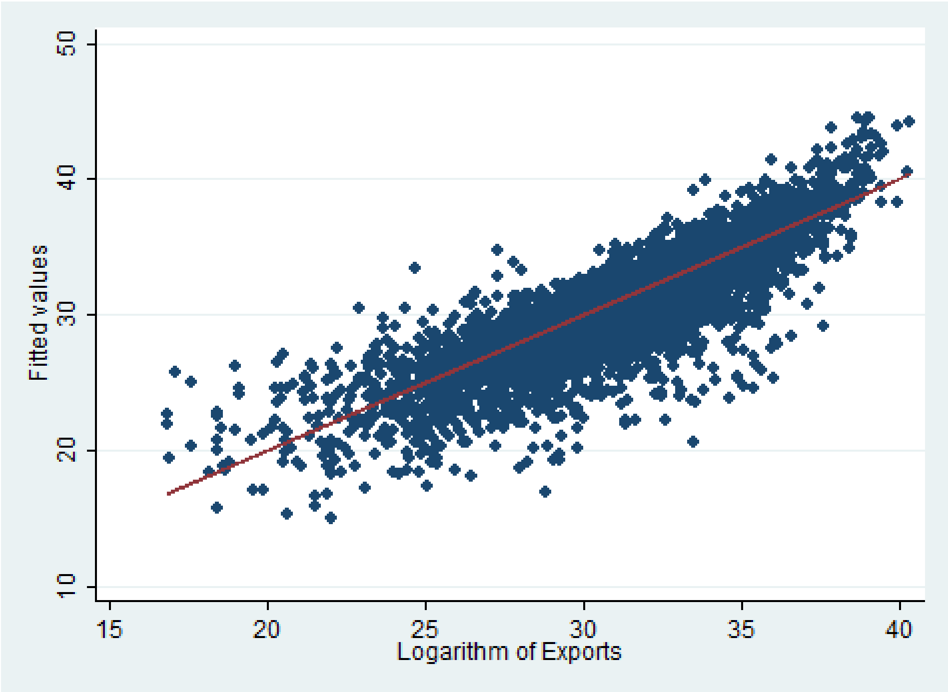


**Fig. 4** OLS (1+X)

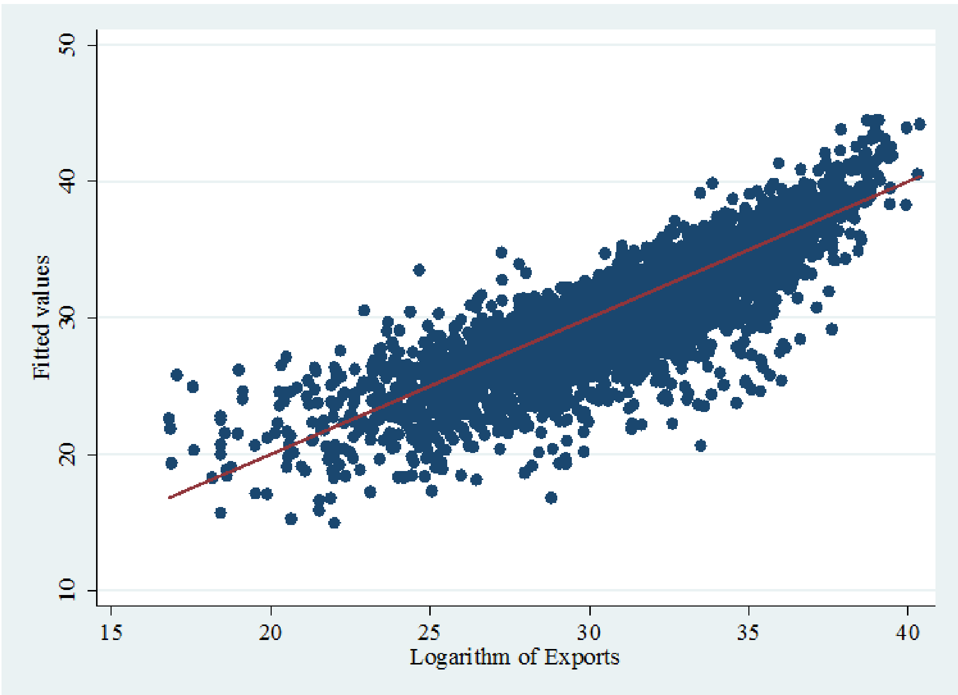


**Fig. 5** Tobit

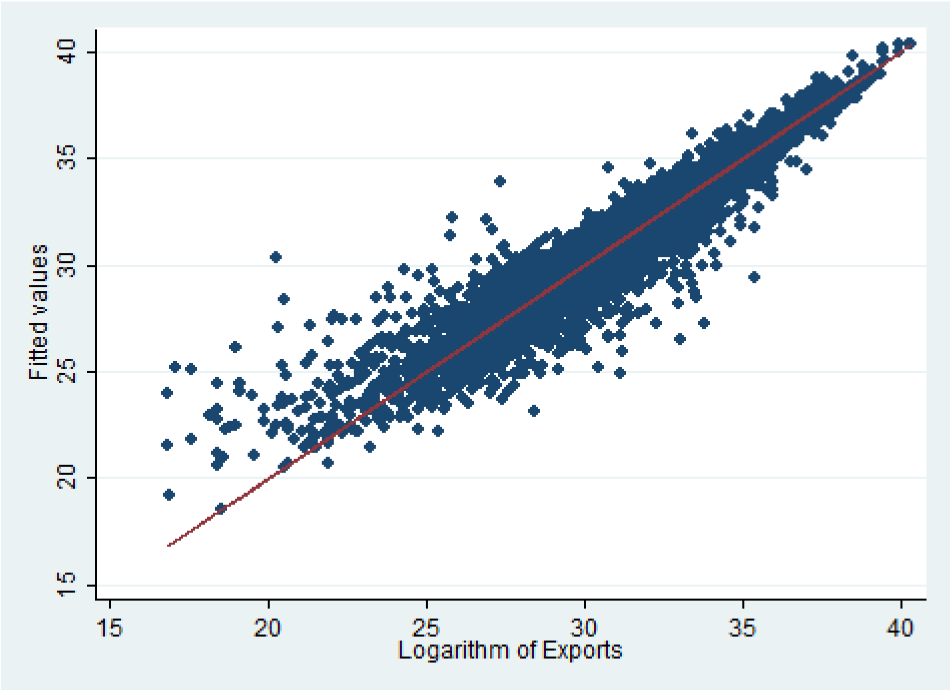


**Fig. 6** Heckman model

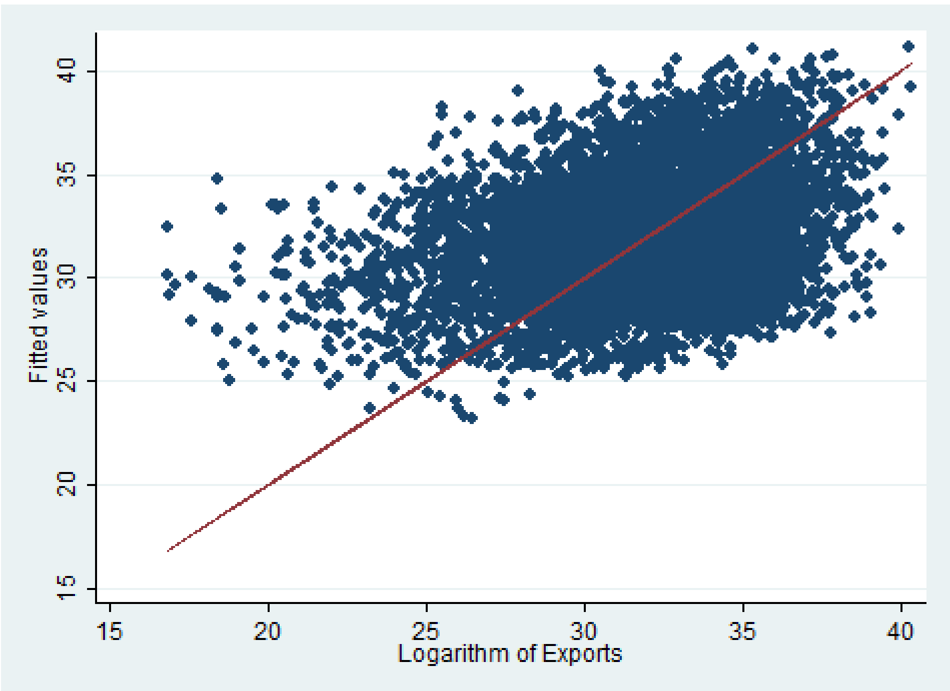


**Fig. 7** Panel fixed effects

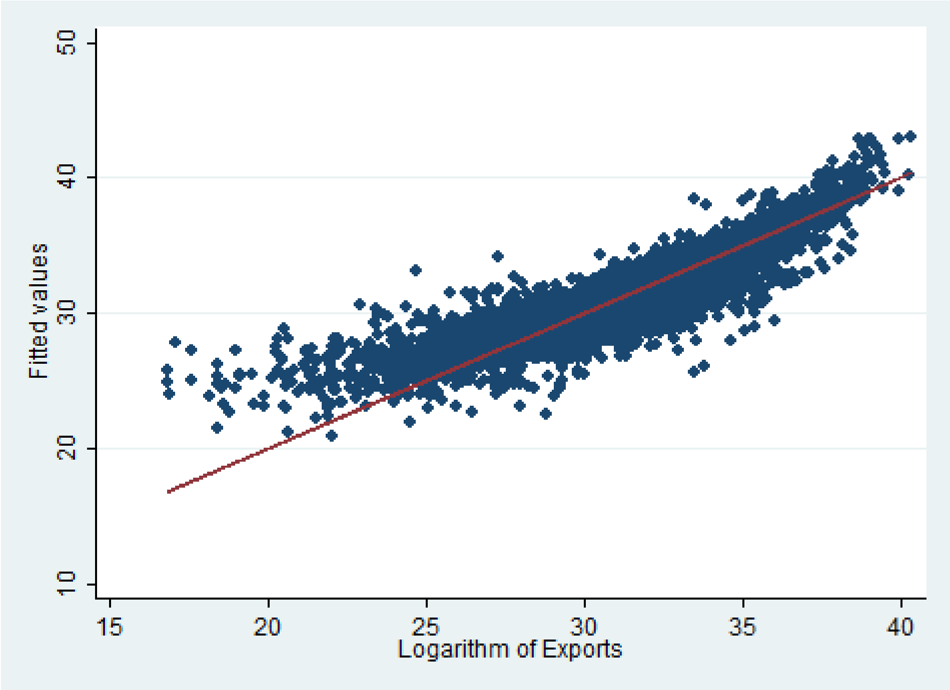


**Fig. 8** Panel random effects

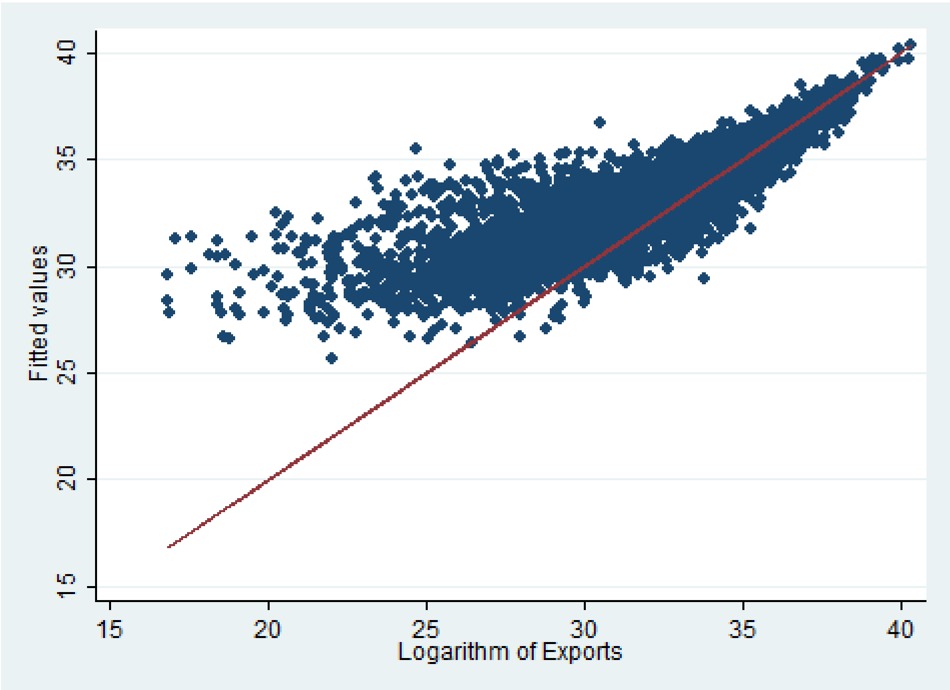


**Fig. 9** Poisson Pseudo Maximum Likelihood

# 7. References


An G, Puttitanun T (2009) Revisiting McCallum's Border Puzzle. Econ Dev Q 23: 167-170

Anderson JE, van Wincoop E (2003) Gravity with gravitas: A solution to the border puzzle. Am Econ Rev 93: 170-192

Anderson JE (1979) A theoretical foundation for the gravity equation. Am Econ Rev 69: 106-116

Anderson JE, Marcouiller D (2002) Insecurity and the pattern of trade: An empirical investigation. Rev Econ Stat 84: 342-352

Andrews M, Schank T, Upward R (2006) Practical fixed effects estimation methods for the three-way error components model. Stata J 6: 461-481

Baier S, Bergstrand J (2009) Bonus vetus OLS: A simple approach for addressing the border-puzzle and other gravity issues. J Int Econ 77: 77-85

Baldwin R, DiNino V (2006) Euros and zeros: The common currency effect on trade in new goods. Center for Economic Policy Research Discussion Papers no 5973

Baldwin R, Taglioni D (2006) Gravity for dummies and dummies for gravity

Baltagi BH, Egger P, Pfaffermayr M (2003) A generalized design for bilateral trade flow models. Econ Lett 80: 391-397

Bergeijk P, Oldersma H (1990) Detente, market-oriented reform and German unification: Potential consequences for the World Trade System. Kyklos 43: 599-609

Bergstrand JH (1985) The gravity equation in international trade: some microeconomic foundations and empirical evidence. Rev Econ Stat 67: 474-481

Bergstrand JH (1989) The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade. Rev Econ Stat 71: 143-153

Bergstrand, JH (1990) The Heckscher-Ohlin-Samuelson model, the Linder hypothesis and the determinants of bilateral intra-industry trade. Econ J 100: 1216-1229

Bikker A, de Vos A (1992) An international trade flow model with zero observations: an extension of the Tobit model. Bruss Econ Rev 135: 379-404

Bouet A, Mishra S, Roy D (2008) Does Africa trade less than it should, and if so, why?: The role of market access and domestic factors. International Food Policy Research Discussion Paper no 770

Brun JF, Carrère C, Guillaumont P, de Melo J (2002) Has distance died? Evidence from a panel gravity model. World Bank Econ Rev 19: 99-120

Burger MJ, van Oort FG, Linders GM (2009) On the specification of the gravity model of trade: zeros, excess zeros and zero-inflated estimation. Spat Econ Analysis 4: 167-190

Bussière M, Schnatz, B (2009) Evaluating China's integration in world trade with a gravity model based benchmark. Open Econ Rev 20: 85-111

Cafiso G (2008) The Euro's influence upon trade: Rose effect versus Border effect. ECB Working Paper 941

Cheng IH, Wall HJ (2005) Controlling for heterogeneity in gravity models of trade and integration. Federal Reserve Bank of St. Louis Rev 87: 49-63

Deardorff A (1998) Determinants of bilateral trade: Does gravity work in a neoclassical world? In: Frankel JA (ed) The Regionalization of the World Economy, 1st edn. University of Chicago Press, Chicago, pp 7-32

De Benedictis L, Vicarelli C (2005) Trade Potentials in Gravity Panel Data Models. Top Econ Anal Policy 5: 1386-1417

Egger P (2000) A Note on the proper econometric specification of the gravity equation. Econ Let 66: 25-31

Egger P, Pfaffermayr M (2003) The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects. Empir Econ 28: 571-580

Egger P, Pfaffermayr M (2004) Distance, trade and FDI: A SUR Hausman-Taylor approach. J Appl Economet 19: 227-246

Egger, P., M. Larch, K. Staub, R. Winkelmann (2011) The Trade Effects of Endogenous Preferential Trade Agreements. American Economic Journal: Economic Policy, forthcoming.

Feenstra RC (2002) Border effects and the gravity equation: consistent methods for estimation. Scot J Polit Econ 49: 491-506

Fidrmuc J (2008) Gravity models in integrated panels. Empir Econ 37: 435-446

Fratianni M, Hoon- Oh CH (2007) On the relationship between RTA expansion and openness. Kelley School of Business, Working Paper no 13, Indiana University

Glick R, Rose AK (2002) Does a currency union affect trade? The time-series evidence. Eur Econ Rev 46: 1125-1151

Haq Z, Meilke K, Cranfield J (2010) Does the gravity model suffer from selection bias? Canadian Agricultural Trade Policy Research Network Working Paper no 90884
Head K, Mayer T (2000) Non-Europe: The magnitude and causes of market fragmentation in the EU. Rev World Econ 127: 284-314
Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47: 153-61
Helliwell JF (1996) Do national borders matter for Quebec's trade? Can J Econ 29: 507-522
Helpman E (1987) Imperfect competition and international trade: evidence from fourteen industrial countries. J Jpn Int Econ 1: 62–81
Helpman E, Melitz M, Rubinstein Y (2008) Estimating trade flows: Trading partners and trading volumes. Q J Econ 123: 441-487
Henderson DJ, Millimet DL (2008) Is gravity linear? J Appl Economet 23: 137-172
Hoon-Oh CH, Selmier II WT (2008) Expanding international trade beyond the RTA border: The case of ASEAN's economic diplomacy. Econ Lett 100: 385-387
Kavallari A, Maas S, Schmitz PM (2008) Explaining German imports of olive oil: evidence from a gravity model. European Association of Agricultural Economists International Congress, August 26-29, 2008, Ghent, Belgium
Leamer E, Stern R (1971) Quantitative international economics. J Int Econ 1: 359-361
Linnemann H (1966) An Econometric study of international trade flows. Dissertation, Netherlands School of Economics
Linders GM, de Groot H (2006) Estimation of the gravity equation in the presence of zero flows. Tinbergen Institute Discussion Papers 06-072/3
Liu X (2009) GATT/WTO promotes trade strongly: Sample selection and model specification. Rev Int Econ 17: 428-446
Manning WG, Mulahy J (2001) Estimating log models: to transform or not to transform? J Health Econ 20: 461-494
Martin W, Pham CS (2008) Estimating the gravity equation when zero trade flows are frequent. MPRA Working Paper 9453, University Library of Munich
Martínez-Zarzoso I, Nowak-Lehmann F, Vollmer S (2007) The log of gravity revisited (in revision). CEGE Discussion Paper 64, University of Göttingen
Mátyás L (1998) The gravity model: Some econometric considerations. World Econ 21: 397-401
Melitz M (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity, Econometrica 71: 1695-1725
Melitz M (2007) North, South and distance in the gravity equation. Eur Econ Rev 51: 971-991
Micco A, Stein E, Ordoñez G (2003) The currency union effect on trade: early evidence from EMU. Econ Policy 18: 315-356
Rose AK, van Wincoop E (2001) National money as a barrier to international trade: The real case for currency union. Am Econ Rev 91: 386-390
Ruiz J, Vilarrubia JM (2007) The wise use of dummies in gravity models: export potentials in the Euromed region. Banco de España Working Papers 0720
Schiavo S (2007) Common currencies and FDI flows. Oxford Econ Pap 59: 536–560
Shepherd B, Wilson JS (2009) Trade facilitation in ASEAN member countries: Measuring progress and assessing priorities. J Asian Econ 20: 367-383
Siliverstovs B, Schumacher D (2009) Estimating gravity equations: to log or not to log? Empir Econ 36: 645-669
Santos Silva JM, Tenreyro S (2006) The Log of Gravity. Rev Econ Stat 88: 641-658
Santos Silva JM, Tenreyro S (2008) Trading partners and trading volume: Implementing the Helpman-Melitz-Rubinstein model empirically. CEPR Discussion Papers 0935
Shepotylo O (2009) Gravity with zeros: estimating trade potential of CIS countries. Kyiv School of Economics Discussion Papers 16
Soloaga I, Winters A (2001) Regionalism in the nineties: what effect on trade? N Am J Econ Financ 12: 1-29
Wang ZK, Winters LA (1991) The trading potential of Eastern Europe. CEPR discussion paper 610
Wei SJ (1996) Intra-national versus international trade: How stubborn are nations in global integration? NBER Working Paper 5531
Westerlund J, Wilhelmsson F (2009) Estimating the gravity model without gravity using panel data. Appl Econ 43: 641-649
Wolf HC (1997) Patterns of intra- and inter-state trade. NBER Working Paper 5939
Yu M (2010) Trade, democracy, and the gravity equation. J Dev Econ 91: 289–300