The Log of Gravity Revisited

By

Inmaculada Martínez-Zarzoso [a]

---

[a] Department of Economics and Ibero-America Institute for Economic Research, University of Goettingen, Germany, and Department of Economics and Instituto de Economía Internacional, Universidad Jaume I, Castellón, Spain. Financial support from the Spanish Ministry of Science and Innovation is gratefully acknowledged (ECO2010-15863). E-mail: martinei@eco.uji.es.

The Log of Gravity Revisited

Abstract

This paper evaluates the performance of alternative estimation methods for gravity models with heteroskedasticity and zero trade values. Both problematic issues, recently addressed by Santos Silva and Tenreyro in an influential paper, are re-examined here. We use Monte Carlo simulations to compare the Pseudo Poisson Maximum Likelihood (PPML) estimator recommended by Santos Silva and Tenreyro, a Gamma pseudo-maximum-likelihood (GPML), a Non-Linear Least Squares (NLS) estimator and a Feasible Generalized Least Squares (FGLS) estimator with more traditional techniques. Additionally, estimates of the gravity equation are obtained for three different data sets with the abovementioned methods. The results of the simulation study indicate that, although the PPML estimator is less affected by heteroskedasticity than others are, its performance is similar, in terms of bias and standard errors, to the FGLS estimator performance, in particular for small samples. GPML presents however the lowest bias and standard errors in the simulations without zero values. The results of the empirical estimations, using three different samples containing real data, indicate that the choice of estimator has to be made for each specific dataset. There is not a general "best" estimator and it is highly recommended to follow a model selection approach using a number of tests to select the more appropriate estimator for any application.

JEL classification: C5, F10

Keywords: Simulations; Poisson regression; constant-elasticity models; heteroskedasticity; Feasible Generalized Least Squares (FGLS); maximum likelihood

1. Introduction

Trade patterns can be extremely well described by the gravity model. The model, in its simplest form, states that bilateral trade between two countries is directly proportional to the product of the countries' income and negatively related to the distance between them. This model has been very successful empirically, mainly due to its high explanatory power. The most common practice in empirical applications has been to transform the multiplicative model by taking natural logarithms and to estimate the obtained log-linear model using Ordinary Least Squares (OLS)[1]. Despite the extensive use of this practice, some econometric problems arise from the prevalence of zero values in trade flows and from heteroskedastic residuals.

Santos Silva and Tenreyro (2006, 2011) propose a Pseudo Poisson Maximum Likelihood (PPML) estimation technique that is consistent in the presence of heteroskedasticity and provides a way of dealing with zero values of the dependent variable. PPML is a special case of the generalized linear model (GLM) framework, in which the variance is assumed to be proportional to the mean. The authors recommend estimating constant-elasticity models using the PPML estimator, instead of applying traditional OLS estimation techniques to the log-linearized models. In their paper, the gravity equation for international trade is taken as an example of constant-elasticity models. Although the general practice has been to estimate the log-linearized version of the gravity model[2], in the related empirical literature, we can also find several attempts to deal separately with heteroskedasticity (e. g., Porojan, 2001) and zero values (e. g., Helpman, Melitz and Rubinstein, 2008).

Martin and Pham (2008) state that the PPML estimator is less subject to bias resulting from heteroskedasticity, but does not prove robust to the joint problems of heteroskedasticity and zero trade flows. Therefore, it could be appropriate for other typical applications of

---

[1] See Oguledo, V.I. and Macphee (1994) for a review of gravity studies using aggregated trade and Martínez-Zarzoso, Pérez-García and Suárez Burguet (2008) for an application using sectoral trade data.
[2] In Chapter 5, Feenstra (2004) presents a revision of articles applying the gravity model for international trade.

multiplicative models with relatively few zero observations, such as the estimation of consumer-demand systems, Cobb-Douglas-type production functions and the STIRPAT (Stochastic Impacts by Regression on Population, Affluence and Technology) model in environmental economics.

In this paper we consider cases with none or with a low percentage of zero values in the dependent variable and focus on the heteroskedasticity problem. Given the extended use of the Feasible Generalized Least Squares (FGLS) estimator to correct for heteroskedastic errors, it seems important to examine and compare its performance with the PPML estimator and with alternative GLM.

The novelties of the present research are threefold. First, we argue that the performance of the PPML estimator should be also compared to FGLS techniques applied to the linearized model and not only to OLS techniques. Second, we consider two different ways of dealing with zero values of the dependent variable; we generate a dependent variable that becomes zero with a certain probability (p) and that remains continuous with probability (1-p), and zeros are randomly created depending on a threshold value of 1 of the regressors to imitate some minimum conditions that are necessary for a positive trade value. Deviating from Santos Silva and Tenreyro (2006, 2011) and Martin and Pham (2008), we will evaluate the performance of the estimators not only by looking at the bias of the estimates, but also by computing their expected loss. Third, estimates of the gravity equation are obtained for three different data sets with two classes of estimators: traditional methods derived from least squares estimators for the log-linearized model and GLM for the unlogged model. For completeness we also estimate a Heckman sample-selection model. In the context of the gravity model of trade, the presence of zeros in the dependent variable is due primarily to the absence of trade, rather than to missing values. This presence raises a problem of selection bias. Heckman-type models are able to tackle the problem of sample selection bias (there are different motives for

participating or not participating in a certain economic activity that have to be identified before running the regression for the agents participating).

In line with Santos Silva and Tenreyro, we use Monte Carlo simulations to evaluate the performance of alternative estimation methods (OLS, FGLS, Nonlinear Least Squares (NLS), Gamma Pseudo Maximum Likelihood (GPML) and PPML) for multiplicative and log models with heteroskedasticity. The results of the simulation study indicate that although the Pseudo Poisson Maximum Likelihood estimator is less affected by heteroskedasticity than other estimators, its performance is similar, in terms of bias and standard errors, to the FGLS estimator performance.

In addition, the results of the empirical estimations, using three different samples containing real data indicate that the choice of estimator has to be made for each specific dataset. There is not a general "best" estimator and it is highly recommended to follow a model selection approach using a number of tests to select the more appropriate estimator for any application.

The next section focuses on a discussion of the PPML and FGLS estimation techniques and their consistency in the presence of heteroskedasticity. Section 3 presents the simulation results in the presence of homoskedasticity and heteroskedasticity, and with a dependent variable that can take on zero values. Section 4 applies the PPML and other alternative estimation techniques to real data and discusses the estimation results. In Section 5, some conclusions are drawn.

2. Alternative Estimators

Let us assume that the original stochastic model is given by

$$y_i = \exp(x_i \beta_i)\varepsilon_i = \mu(x\beta)\varepsilon_i, \tag{1}$$

where $E(\varepsilon_i \mid x) = 1$.

Assuming that y_i is positive, the model can be linearized by taking logs:

$$\ln y_i = x_i \beta_i + \ln \varepsilon_i, \tag{2}$$

where $\ln E(\varepsilon_i \mid x) = 0$; $E(\ln \varepsilon_i \mid x) \neq 0$.

The original multiplicative model given by Equation 1 can be directly estimated using NLS or maximum likelihood techniques, whereas Model 2 can be estimated simply by OLS. The inequality $E(\ln \varepsilon_i \mid x) \neq 0$ is known as Jensen's inequality. In the presence of heteroskedasticity, Least Squares (LS) estimation is no longer efficient. This problem can be resolved by controlling for heteroskedasticity.

The estimation technique proposed by Santos Silva and Tenreyro (2006) is a pseudo maximum likelihood estimator based on some assumptions about the functional form of the conditional variance. Under the assumption that the conditional variance is proportional to the conditional mean, β can be estimated by solving a set of first-order conditions:

$$\sum_{i=1}^{n} \left[y_i - \exp\left(x_i \widetilde{\beta} \right) \right] x_i = 0, \tag{3}$$

where y_i is the dependent variable, x_i are the explanatory variables, and β are the parameters to be estimated. The estimator based on Equation 3 gives the same weight to all observations, whereas the LS and NLS estimators give more weight to observations with large $\exp(x_i \beta)$.

The data do not have to follow the Poisson distribution and the dependent variable does not have to be an integer. Since the assumption of proportionality between the conditional variance and the conditional mean does not always hold, inference should instead be based on a robust covariance matrix estimator that specifically corrects for heteroskedasticity in the model.

Within the GLM framework it is common to consider a more general form of the variance function given by

$$v(y \mid x) = k(\mu(x\beta))^{\lambda} \tag{4}$$

where λ is finite and non-negative and its value determines different GLM families. We obtain the NLS estimator when $\lambda=0$. We obtain the Poisson class when $\lambda=1$. We obtain the

gamma, the homoskedastic log-normal, the Weibull and the chi square in the case λ=2 with the appropriate specification of a distribution. Finally, when λ=3, the Inverse-Gaussian family is obtained.

As mentioned above, the traditional way to estimate constant-elasticity models is to linearize the original multiplicative model using a log-log transformation. The rational for using a log-log model can come from a desire to generate an estimate that easily yields elasticities or from a need to deal with dependent variables that are skewed to the right. Log transformation also improves precision and reduces the influence of outliers on the estimates. However, since no one is interested in log model results in the log scale per se, those results have to be retransformed to the original scale in order to be able to quantify the average response to a covariate (Manning, 1998).

Within the class of least squares estimators, the FGLS estimator is efficient in the presence of heteroskedasticity and can be applied to the linearized model. The FGLS estimator weighs the observations according to the square root of their variances[3], and is given by

$$\beta_{FGLS} = \left(x'\hat{\Omega}^{-1}x\right)^{-1}x'\hat{\Omega}^{-1}y, \tag{5}$$

where $\hat{\Omega}$ is the weighting matrix. $\hat{\Omega}$ is estimated using the OLS squares residuals ($\hat{u}_i$), to obtain an estimate for the variance of $\hat{u}_i$ ($\widehat{\text{var}}\ \hat{u}_i$). Then, $\hat{\Omega} = 1/\sqrt{\widehat{\text{var}}\ \hat{u}_i}$.

It can be shown that even if the weights used in FGLS estimation are biased (resulting in a biased estimation of the residuals variance), FGLS would still provide consistent estimates when the appropriate retransformation technique is used[4]. In case of an unknown form of heteroskedasticity, FGLS can be applied and the variance of the disturbances must be estimated. This method should be well suited to estimating regression coefficients in the

---

[3] Heteroskedasticity is assumed to be alinear function of the explanatory variables, as in Wooldridge (2009), pages 282-283.
[4] With transformhed dependent variables, the mean response is the mean of the retransformed estimate of y, which depends not only on the mean of the x's, but also on their distribution (Manning, 1998).

presence of heteroskedasticity. Henceforth, the comparison should be made between FGLS-like estimators and GPML, NLS or PPML estimation.

The main difficulty encountered when dealing with log dependent variables is that log scale results are of little interest. Most economic policy decisions are taken based on monetary values requiring therefore retransformation of the log scaled predictions. Retransformation does not present any problem when errors are normal and homoskedastic, however, retransformation biases arise when one of these do not hold. A common practice in health economics has been to switch to GLM from log-transformed models when errors are heteroskedastic. However, although GLM can deal with heteroskedastic errors and do not lead to retransformation bias, OLS/FGLS of log-transformed variables could be more efficient in some cases (Manning and Mullahy, 2001). GLM can suffer substantial precision losses if the log-scale residuals have high kurtosis or if the variance function is misspecified[5]. The key question that remains is how to determine which method is best for specific empirical applications. Manning and Mullahy (2001) developed a method for determining which estimation method to choose for any empirical application which is based on using a number of tests that are easy to implement.

3. Simulation Study

A data set with the same properties as described by Santos Silva and Tenreyro (2006) is generated to compare different estimation techniques in the presence of heteroskedasticity. We extend the procedures used by the authors in two ways. First, we add an alternative estimation technique, the FGLS estimator, which is applied to the linearized model to correct for heteroskedasticity. Second, instead of using a rounding approach to generate zeros in the dependent variable[6], two alternative methods are proposed; one consists of setting

---

[5] However, estimates are still consistent.
[6] As done in Santos Silva and Tenreyro (2006).

observations to zero randomly and the other consists of setting observations to zero for particular groups.

The multiplicative constant elasticity model considered is

$$y_i = \mu(x_i\beta)\eta_i = (\widetilde{\beta}_0 \, X_{1i}^{\beta_1} X_{2i}^{\beta_2})\eta_i \,, \tag{6}$$

η_i is a log-normal random variable with mean 1 and variance σ_i^2, in the absence of heteroskedasticity. x_{1i} is log-normal and x_{2i} is dichotomous with values of e^0 and e^1 with a probability of 0.6 for the first value and a probability of 0.4 for the second value.

Equation 5 can also be expressed as

$$\mu(x_i\beta) := E[y_i \mid x] = \exp(\ln \widetilde{\beta}_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i}) \,, \tag{7}$$

where $X_{1i} := \ln x_{1i}$; $X_{2i} := \ln x_{2i}$ and $\ln \widetilde{\beta}_0 := \beta_0$.

From distributional theory, it follows that X_{1i} is drawn from a standard normal distribution and X_{2i} is a binary variable that takes the value 1 with a probability of 0.4. The two covariates are independent and the true values for the coefficients are: $\beta_0 = 0$, $\beta_1 = 1$ and $\beta_2 = 1$.

In line with Santos Silva and Tenreyro (2006), we assess the performance of different estimators (NLS, GPML, PPML, OLS and FGLS)[7] under homoskedasticity and heteroskedasticity. We can distinguish four cases:

Case 1: $V(\eta_i) = \mu(x_i, \beta)^{-2}; V[y_i \mid x] = 1$

Case 2: $V(\eta_i) = \mu(x_i, \beta)^{-1}; V[y_i \mid x] = \mu(x_i, \beta)$

Case 3: $V(\eta_i) = 1; V[y_i \mid x] = \mu(x_i, \beta)^2$

Case 4: $V(\eta_i) = \exp(x_{2i}) + \mu(x_i, \beta)^{-1}; V[y_i \mid x] = \mu(x_i, \beta) + \exp(x_{2i})\mu(x_i, \beta)^2$

In Case 1, the NLS assumptions hold. In Case 2, the conditional variance equals its conditional mean, as in the PPML assumptions. OLS, FGLS and GPML conditions are fulfilled in Case 3 (homoskedasticity). And finally, in Case 4, the conditional variance does

---

[7] Tobit and truncated OLS simulation results are not presented since these methods show a very poor performance.

not only depend on the mean, but is also a function of one of the explanatory variables (heteroskedasticity). We argue that in Case 4 of heteroskedasticity, which was identified as a severe problem in many applications (e.g., when analyzing trade flows in the framework of the gravity model), FGLS, rather than OLS, should be used for comparison purposes.

In a first set of simulations, we study the performance of the estimators for the different models considered and for the four specifications of the disturbances outlined above. We focus on the estimators with a good performance in the experiments of Santos Silva and Tenreyro (2006), add the FGLS estimator and disregard those with a poor performance (Tobit, OLS (y+1), truncated-OLS).

The truncated and the censored OLS models (in logarithms) produce results that are inferior to those obtained from the classical OLS model (that disregards the observation with zero values) in terms of bias and standard errors. This result is consistent with Martin and Pham (2008).

Initially, we also considered the Tobit estimator which applies to situations in which outcomes can only be observed over some range, either because actual outcomes cannot reflect desired outcomes or because of measurement inaccuracy (rounding). However, the gravity equation only predicts zero trade if incomes of one or both countries are zero; therefore, desired outcomes are always positive. In addition, trade values are reported in COMTRADE even for very small amounts (up to $1). Therefore, zero flows seem to result from a binary decision-making process, rather than from censoring. This is also in accordance with the Helpman et al. (2008) model.

Deviating from Santos Silva and Tenreyro (2006) and Martin and Pham (2008), we will evaluate the performance of the estimators not only by looking at the bias of the estimates, but also their expected loss. Most economic simulation studies consider unbiasedness or a small bias to be the most desirable property of an estimator. Therefore, the bias is used as the main criterion to compare the quality of different estimators. But this approach could be misleading

in some cases since –due to the fact that over- and under-estimations cancel each other out– unbiased estimators are not necessarily good estimators[8]. A feature of our results for some of the considered estimators is the wide variation in the direction of the bias in parameter estimates; for example, the bias in the estimate of the slope coefficients resulting from use of the PPML estimator is not consistently negative. For this reason, although the bias is considered to be very low in some of the four cases, the expected loss is not. By considering the expected loss, the direction of the bias is also taken into account in evaluating the performance of alternative estimators.

In statistical decision theory, one therefore looks at the risk of an estimator, defined as its expected loss:

$$R(\beta, \hat{\beta}) = E_\beta L(\beta, \hat{\beta}). \tag{8}$$

The way the loss function $L(\beta, \hat{\beta})$ is defined depends on the individual needs in each statistical analysis. As argued above, we only consider the class of loss functions where over- and underestimations cannot cancel out, as in case of the bias. The simplest loss function which is adequate for our purposes is the absolute error loss:

$$L(\beta, \hat{\beta}) = \left| \beta - \hat{\beta} \right|. \tag{9}$$

As an alternative loss function, the squared error loss could also be used.

In a second set of experiments, we study how zero values in the dependent variable affect the performance of the estimators, using a data-generation mechanism that produces a percentage of observations with zero values in y_i, but that differs from those used by Santos Silva and Tenreyro (2006)[9] and Martin and Pham (2008). In Santos Silva and Tenreyro (2006), the zeros in the dependent variable are generated by rounding to the nearest integer the values of

---

[8] We give a simple example to illustrate this: let the true parameter be $\beta = 1$. We define an estimator $\hat{\beta} = 0$ half of the time and $\hat{\beta} = 2$ the other half of the time. Clearly, this is a very bad estimator, but it is unbiased. The constant estimator $\hat{\beta} = 1.1$, in contrast, is biased but obviously better than the above estimator.

[9] In Santos Silva and Teneyro (2006), the dependent variable was generated by rounding the values of y_i to the nearest integer.

y_i obtained in the first set of simulations. The authors mention that zeros could be a result of rounding errors or of missing observations that are wrongly recorded as zeros. As they recognize, this procedure will bias the estimates and the purpose of the experiment is just to measure the magnitude of the bias. To check the robustness of the results we suggest a slightly different way to investigate cases of zero trade or related problems, namely setting a certain percentage of the dependent variable to zero. We do this in two different ways: first, we choose a certain percentage, say 15 percent[10], of all observations and set it to zero, and second, we only create zeros in a certain group, assuming that there is an underlying pattern that is responsible for this, say half of the 30 percent poorest/smallest countries, according to one or more of the independent variables. This pattern is reasonable for many relevant real cases, including the gravity model, since zero trade mainly occurs among poor or small countries.

Tables 1 and 2 summarize simulation results for sample sizes of 1,000 and 10,000 replications. The expected loss, the bias and the standard error of the two parameters of interest (β_1 and β_2) are presented. Table 1 shows the results for the first set of simulations. The NLS shows a lower bias and expected loss than the other estimators only in Case 1. In Cases 3 and 4, the expected loss and bias of the NLS estimator are the highest. The PPML is the best in Case 2, in which the underlying error structure is consistent with the PPML estimator. In Cases 3 and 4, the performance of FGLS and PPML are comparable. FGLS estimates display lower expected loss, biases and standard errors than PPML in Case 3. The bias in the estimate of β_1 is only -0.0003 for FGLS and -0.0027 for PPML. For β_2 the corresponding biases are 0.0001 and -0.0008. In Case 4, FGLS estimates also present lower expected loss and standard errors for β_1 and β_2, but PPML shows slightly lower biases.

Table 1. Simulation results, regular case

---

[10] Bilateral trade among developed countries presents this pattern. Around 10 to 15 percent of the bilateral flows are zeros.

Table 2 shows the results for the second set of simulations, in which the dependent variable is generated with a percentage of zero values (15 percent). The NLS estimator yields estimates that are strongly biased in Cases 3 and 4. Traditional OLS estimates also present large biases and expected losses in Cases 1, 2 and 4. However, correcting for heteroskedasticity using FGLS considerably reduces the expected losses and biases in Cases 1, 2 and 4. When the zero values are generated randomly, the Gamma and the FGLS perform quite well in Cases 3 and 4, and both present lower expected losses and standard errors than the PPML estimation. However, when the zero values are generated with a given pattern, the Gamma estimator yield estimates that are strongly biased in all cases, especially for β_1, which bias is around 25 percent in all four cases. FGLS shows lower expected losses and standard errors than PPML in Cases 2, 3 and 4. Also in terms of bias, the bias in the estimate of β_1 for FGLS is always lower than for PPML and it is lower than 1 percent in Cases 2 and 3 and around 5 percent for Cases 1 and 4.

Table 2. Simulation results with zeros

Additionally, we repeated the simulations with higher percentages of zero values (40 percent and 60 percent). The results for FGLS and PPML are shown in the appendix. A similar pattern persists for the FGLS estimator, with β_1 biases always lower than 2 percent for Cases 2 and 3, and around 7 percent and 5 percent, respectively, for Cases 1 and 4. The β_2 biases are consistently lower than 1 percent for Cases 3 and 4, and around 5 percent and 8 percent, respectively, for Cases 1 and 4. The biases of β_1 and β_2 for PPML increase with higher percentages of zero values[11], and more rapidly than for FGLS. The expected losses are around 10 percent for β_1 and around 15 percent for β_2 in Cases 3 and 4, with 40 percent being zero values.

---

[11] Martin and Pham (2008), using different data-generating processes, also show that the PPML estimator is highly vulnerable to the presence of a high percentage of zero values in the dependent variable. However, these results have been challenged by Santos Silva and Tenreyro (2011).

As a robustness check, and as suggested by Santos Silva (personal communication), we increased the sample size to 10,000 observations, which amount is consistent with some of the cross-sectional analysis in the framework of the gravity equation. Whereas the bias of the PPML decreases, the FGLS bias remains almost constant. The standard errors of the PPML estimator is considerably lower, but still twice the FGLS standard errors. Therefore, using PPML when the sample is large and when there is measurement error in the dependent variable could be recommendable. However, for small samples, FGLS could be a perfect choice to deal with the problem of heteroskedastic disturbances.

4. The Gravity Model: Estimation Results

4.1 Original versus log-log version of the gravity model

According to the generalized gravity model of trade, the volume of exports between pairs of countries in year t, X_{ijt} is a function of their incomes (Y_{it}, Y_{jt}), their incomes per capita (YH_{it}, YH_{jt}), their geographical distance ($DIST_{ij}$) and a set of dummies that represents any other factors aiding or preventing trade between pairs of countries (F_{ij}):

$$X_{ijt} = \beta_0\, Y_{it}^{\beta_1} Y_{jt}^{\beta_2} YH_{it}^{\beta_3} YH_{jt}^{\beta_4} DIST_{ij}^{\beta_5} F_{ij}^{\beta_6}\, u_{ijt}\,, \tag{10}$$

where u_{ijt} is the error term.

Equation (10) in its multiplicative form can be directly estimated using GLM. GLM assumes that there is a function that explains the relationship between the variance and the mean (see equation 4 above) and does not assume constant variance.

However, the most prevalent modelling approach is to use OLS or similar techniques to the log-linearized model. Taking natural logarithms from Equation 10 and replacing F_{ij} for a number of specific variables, the log-log model is given by

$$\begin{aligned} LX_{ijt} &= \alpha_{ij} + \beta_1 LY_{it} + \beta_2 LY_{jt} + \beta_3 LYH_{it} + \beta_4 LYH_{jt} + \\ &+ \beta_5 LDIST_{ij} + \beta_6 Border_{ij} + \beta_7 Comlang_{ij} + \beta_8 Z_{ijt} + \sum_k \lambda_k FTA_{ij} + \mu_{ijt} \end{aligned} \tag{11}$$

where:

L denotes variables in natural logs,

X_{ijt} are the exports from country *i* to country *j* in period *t* in current US$,

Y_{it}, Y_{jt} indicate the GDP of countries *i* and *j,* respectively, in period *t* in current PPP US$*,*

YH_{it}, YH_{jt} denote incomes per capita of countries *i* and *j*, respectively, in current PPP US$ per thousand inhabitants in period *t,*

$DIST_{ij}$ is the Great Circle distance between countries *i* and *j*.

Border is a dummy that takes the value of 1 when countries *i* and *j* share a border, and

Comlang is a dummy that takes the value of 1 when countries *i* and *j* speak the same language.

Z are additional variables that could foster or prevent trade between pairs of countries and could be, in some cases, time variant, for example, a dummy that takes the value of 1 if the trading countries share a common currency, or a dummy that takes the value of 1 if they had a colonial relationship in the past, 0 otherwise.

The model includes trading blocs' dummy variables, defined as $FTA_{ij,}$ which evaluate the effects of free trading agreements (FTAs). Integration dummies are described in more detail below. α_{ij} are the specific effects associated with each bilateral trade flow. They are a control for all the omitted variables that are specific for each trade flow and that are time invariant.

A high level of income in the exporting country indicates a high level of production, which increases the availability of goods for export. Therefore, we expect β_{1} to be positive. The coefficient of Y_{j}, β_{2}, is also expected to be positive since a high level of income in the importing country attracts higher imports. The coefficient for income per capita of the exporters, β_{3}, may be negatively or positively signed depending on the type of goods exported. The coefficient of the importer income per capita, β_{4}, also has an ambiguous sign, for similar reasons. Another factor that may influence these coefficient estimates is the composition effect that influences supply and demand. Each country produces and exports a different mix of commodities (supply) and the mix of goods demanded is also different for

each country. The distance coefficient is expected to be negative since it is a proxy of all possible trade costs.

In some of the results presented below, specific trading bloc dummies are included to model trade-creation and trade-diversion effects of trade (see Soloaga and Winters (2001); Chen and Tsai (2005), Carrère (2005) and Martínez-Zarzoso, Nowak-Lehmann and Horsewood (2009)). The specification of the Viner's trade creation and trade diversion for a single period is given by

$$LX_{ij} = EV_{ij} + \sum_k \gamma_k D_k + \sum_k \delta_k D_{ki} + \sum_k \rho_k D_{kj} \quad , \tag{12}$$

where LX_{ij} denotes exports from country i to country j *in natural logarithms*, and EV_{ij} is defined as the rest of explanatory variables of the gravity equation above. D_k is a dummy that takes the value 1 if both countries, i and j, belong to the same economic bloc, 0 otherwise. D_{ki} is a dummy that takes the value 1 if i is a member of bloc k and j belongs to the rest of the world, 0 otherwise. D_{kj} is a dummy that takes the value 1 if j is a member of bloc k and i belongs to the rest of the world, 0 otherwise. γ_k measures the extent to which trade exceeds normal levels if both countries, i and j, are members of the bloc, δ_k measures the extent to which members' exports levels to non-member countries are higher than normal, and ρ_k measures the extent to which members' import levels from non-member countries are higher than normal. δ_k and ρ_k could be interpreted as a measure for trade-diversion effects, but they might also combine trade-diversion and openness effects.

4.2 Main results

Different versions of Models 10 and 11 are estimated using two classes of estimators: least squares for models with logged dependent variables, and GLM with a logarithmic function.

Within the first class we consider four estimators: OLS, FGLS, Harvey model[12] and the Heckman selection model. Within the second class, we use tree estimators: PPML, GPML and NLS. Observations with zero values are simply dropped when estimate the OLS, FGLS and Harvey models. The FGLS estimator is computed two steps. The squared OLS residuals are computed (e^2) and regressed upon the gravity-model explanatory variables and a constant. Then $1/\sqrt{\hat{e}^2}$ are the weights used in the second step weighted least squares regression.

The gravity model is estimated for three different data sets. The first one is a sample of 180 countries over the period from 1980 to 2000 from Rose (2005). The second data set consists of a sample of 65 countries from Márquez, Martínez and Suárez (2007), with data for every five years over the period from 1980 to 1999. Finally, a third sample of 47 countries from Martínez-Zarzoso et al. (2009) is used, covering the period from 1980 to 1999. The percentage of zero export values in the three samples is 13, 15, and 23 percent, respectively. The traditional gravity model, as well as the theoretically justified gravity model (Anderson and van Wincoop, 2003) with multilateral resistance terms specified as exporter and importer dummy variables, are considered.

The distribution of the dependent variable, total exports, in the three samples shows right-skewness since there are no negative exports and some pair of countries do not trade to each other. The summary statistics for exports in the first sample for the year 1990 are mean= 1.82e+08, variance= 3.29e+18, skewness= 28.59645, and kurtosis= 1155.022. An attempt to make the distribution more normal consists on taking natural logarithms. The equivalent summary statistics for log exports are mean= 14.17866, variance= 16.51199, skewness= -.8381542, and kurtosis= 5.054987. Although transformation makes the distribution closer to a normal, the usual test still rejects the normality assumption.

---

[12] Harvey's model of multiplicative heteroskedasticity (Harvey, 1976) has also been estimated since it is a very flexible and model that includes most of the useful formulations as special cases. The general formulation is $\sigma_i^2 = \sigma^2 \exp(z_i' \alpha)$.

In the other two samples with 65 and 47 countries respectively, whereas unlogged exports are badly skewed to the right, the log of exports is approximately normal for 1990 with skewness close to zero and kurtosis close to 3. According to Manning (1998), in such cases estimates based on logged models are usually more precise and robust than direct analysis of the unlogged original dependent variable. Retransformation presents no problem when errors are linear, normal and homoskedastic. One of the main reasons to switching to GLM from log-transformed models in applied research has been due to heteroskedasticity of unknown form. But if errors are normally distributed, a transformation can still be applied that accounts for any form of heteroskedasticity (Baser, 2007). We applied the method proposed by Baser (2007) to calculate the retransformed expected value of exports.

Table 3 shows the results for the traditional gravity model, as specified in Equations 9 and 10 for the first data set[13]. With comparative purposes, we show the results for the year 1990, as in Santos Silva and Tenreyro (2006) (SST). However, the estimated parameters are not directly comparable since the number of countries is different; as is the definition of the dependent and some of the independent variables. Our purpose is only to present estimation results for different data sets in order to evaluate/put into perspective the use of the PPML as the proposed "new workhorse" for the estimation of gravity models[14].

Table 3. Estimation results: traditional gravity model

The estimated coefficients for the income variables are closer to the theoretical value of 1 when the equation is estimated using OLS and FGLS, Harvey and Heckman. The fact that the income elasticities for the exporter and importer countries are equal in magnitude, according to the PPML estimation, is not indicative of a correct specification. Asymmetries can also be theoretically justified. The standard errors are significantly lower for the least squares estimators and therefore the estimates are more accurate.

---

[13] The theoretically modified gravity model was also estimated for this sample using OLS and FGLS techniques, but the PPML estimation presented some problems since the maximization algorism could not find a solution.
[14] SST page 649.

In contrast to SST (2006), the variable *common border* is also statistically significant when PPLM is used, compared to OLS and FGLS estimates. The OLS and FGLS estimates for *FTA,* indicates that the integration dummy is statistically significant, whereas using PPML, it is non-significant and much lower in magnitude than when using least squares methods. Instead, the GPML model produces a significant coefficient for the FTA dummy. Concerning geographical distance, in line with SST the estimated elasticity is significantly lower in magnitude in the PPLM estimation, whereas in the FGLS and Harvey model, it is lower than in OLS but higher than in PPML.

Table 4 shows the results for the theoretically justified gravity model for the second and third data sets. The first three columns show the estimated coefficients for the OLS, FGLS and Harvey models. Column 4 shows the Heckman selection model results that control for zero values for the dependent variable. The last three columns present the results obtained using GLM (PPML, GPLM and NLS).

Comparing the results of the different specifications, we observe for both samples that the role of distance is significantly reduced under Poisson and NLS. Trading countries having a common language always have a positive effect on trade, but the effect is also substantially reduced (halved) when using Poisson in comparison to OLS, FGLS, Harvey or Heckman. A common border consistently has a positive effect on exports in the first sample (65 countries) under all estimation techniques. However, for the second sample (47 countries) OLS and FGLS and Harvey predict no significant effects and Poisson produces a positive and significant effect. The coefficient is not shown in the fourth column because the variable common border was only used in the selection equation.

With respect to the free-trade agreement variables, for the first sample only PPML and NLS predict a significant and positive effect. For the second sample, PPML predicts always positive effects but the estimate is not significant for the EU, whereas the additional techniques produce negative effects for the EU agreement and positive and significant effects

for the NAFTA, CACM and CARICOM agreements. However, as stated by some authors, the estimation of gravity models using cross-sectional data does not provide a good way to identify the effects of free trade agreements (Baier and Bergstrand, 2007).

Table 4. Estimation results: Theoretically justified gravity model

A panel data version of the gravity model is also estimated, adding to the traditional specification dyadic random effects and time effects. The results are shown in Table 5. The first column presents the result for the random effects model assuming homoskedasticity and uncorrelated errors. The second column allows for heteroskedastic errors and the third assumes autocorrelation of first order. Columns four to seven show the GLM estimates: PPLM, GPLM and NLS respectively. The log-linear models (columns 1 to 3) show more accurate coefficients than the GLM.

Table 5. Estimation results for panel data (1980-1999)

Since many other issues come into play when estimated gravity equations with panel data, as for example what is the correct way to control for multilateral resistance, we refer to the results obtained in Martínez-Zarzoso et al. (2008) for the 47 country sample.

4.3 Robustness check

In this section, we first present the results obtained when testing for the pattern of heteroskedasticity. In addition, we investigate the out-of-sample forecast performance of the alternative estimates.

To determine if the pattern of heteroskedasticity assumed by the different models is acceptable, a Park-type test is used to check for the adequacy of the log-linear model versus alternative GLM. The Park-type test can be directly derived from Equation (4) in Section 2 above. The test consists of estimating the equation specified as

$$\ln(y_i - \hat{y}_i)^2 = \ln(\lambda_0) + \lambda_1 \ln \hat{y}_i + \varepsilon_i .$$ (13)

Acceptance of the null hypothesis: H_0: $\lambda_1=2$, based on a robust covariance estimator, would be in support of the log-linear model. The results of the test performed for the 65 countries sample are presented in the first part of Table 6. The point estimate for λ_1 in 1990 is 1.829 and the estimates for other years are very similar. Below the point estimates, the 95 confidence intervals are also reported in Table 6. Although those intervals do not contain the value 2, they are closer to this value than to value 1, that will be in support of the PPML assumption (the variance is proportional to the mean).

Since log-linerization of equation (4) is valid only under restrictive conditions of the conditional distribution of y_i, we also used a robust alternative. As suggested by Manning and Mullahy (2001) and Deb, Manning and Norton (2008) we have also estimated the modified version of the Park test as a GLM with log link where the dependent variable is $(y_i - \hat{y}_i)^2$ and the explanatory variable is $x\hat{\beta}$ from each one of the GLM estimators of y on x, using a robust variance covariance matrix. The second part of Table 6 shows the results for several years using Rose data. Among the GLM estimators we have chosen the Gamma family since it presents the lowest AIC and BIC values. The results are presented in the second part of Table 6. The point estimate for λ_1 in 1990 is 1.586 and the estimates for other years are slightly higher. Below the point estimates, the 95 confidence intervals are also reported. According to those results, the variance always exceeds the mean and a Gamma family (within the GLM models) would be more appropriate than a Poisson family. We also estimated $x\hat{\beta}$ using the Poisson and Gaussian families, and the confidence intervals did not contain 1 and where always closer to 2. For the 65 country sample the results for λ_1 lie on the 95% confidence interval (1.66-1.75) and for the 47 country sample, the estimated confidence interval for λ_1 is (1.77-1.85). According to these results the more appropriate GLM family could be the gamma.

SST used instead the results from a Gauss-Newton regression to test for the proportionality assumption (the conditional variance is proportional to the conditional mean). The Gauss-Newton regression is given by

$$(y_i - \hat{y}_i)^2 \big/ \sqrt{\hat{y}_i} = \lambda_0 \sqrt{\hat{y}_i} + \lambda_0 (\lambda_1 - 1) \ln(\hat{y}_i) \sqrt{\hat{y}_i} + \mu_i. \tag{14}$$

The proportionality hypothesis can be checked by testing for the significance of $\lambda_0 (\lambda_1 - 1)$ in Equation 14. The hypothesis will be accepted if the parameter is statistically insignificant. Equation 14 is derived using a Taylor series expansion around $\lambda_1 = 1$ (which is an approximation) and assuming that the variance is proportional to the mean to choose the appropriate weights in equation (14). We consider that the approach proposed by Manning and Mullahy (2001) and Deb, Manning and Norton (2008) is more straightforward and gives a confidence interval for λ_1 which is more useful to choose the more appropriate GLM estimator.

As pointed out by SST, the proportionality assumption is unlikely to hold; therefore, the PPLM does not fully account for the heteroskedasticity in the model. The authors mentioned that the inference should then be based on a robust covariance matrix estimator. However, another alternative is to specify the variance in a more accurate way: using a gamma family if λ_1 is close to 2, for example.

In this context (with λ_1 close to 2) FGLS could also be a good method. The shortcoming of standard FGLS techniques is that the relationship between the residuals and the independent variables is restricted to a linear model. In the presence of Jensen´s inequality, as well as for other functional forms different from the linear model, this simplifying assumption lowers estimation quality. Nevertheless, for most practical purposes, this problem could be solved by using more sophisticated methods based on the results obtained in Carroll (1982), Robinson (1987) and Delgado (1992 and 1993). These authors propose a semi-parametric version of FGLS that consists of estimating the relationship between the residuals and the independent variables, allowing for an unknown functional form and using non-parametric techniques.

The performance of the log-linearized models can be compared using the root mean squared error (RMSE) reported in Tables 3 and 4 (in the first and second half of the table). Among those models, the estimates of the FGLS model present the lowest RMSE. In order to use the RMSE to compare log-log models with GLM models[15], we have re-calculated the RMSE for log-log models using the retransformed predicted values (retransformation proposed by Baser, 2007). The results presented in Tables 3 and 4 (RMSE retrans) indicate that NLS presents the lowest value.

In order to compare different GLM models, the Deviance and Dispersion of the different families can also be compared (See Tables 3 and 4, left hand side). The large Dispersion shown by the Poisson estimates is another indicator that the Poisson distribution is not a good choice. Since the distribution of exports presents signs of overdispersion (the variance is larger than the mean), we scaled the standard errors using the square root of the Pearson chi-square dispersion. The coefficients were identical to the previous analysis but the standard errors were significantly higher in magnitude. The GPLM present (in all cases) the smaller deviance and dispersion out of the three GML methods considered.

SST show the results obtained from a Ramsey-RESET test of specification error of functional form that supported the PPML estimator. We also computed it, and the PPML estimator did not pass the test. None of the models estimated for a single year pass the RESET test, apart from the probit model estimated in the first step of the Heckman model. These results could be due to the fact that relevant variables have been omitted. This problem is solved by using panel data since the introduction of exporter-and-year and importer-and year fixed effects in addition to dyadic fixed effects is an accepted way to control for the omission of relevant variables in gravity models (Baldwin and Talglioni, 2006; Martínez-Zaroso et al., 2009).

We also computed the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Although the results obtained for log-linear models and for GLM are not

---

[15] Following Joao Santos Silva recommendation we retransform the predicted values obtained from log-.log models to make a valid comparison.

comparable, we can use those two criteria to compare models within the first and second class. Whereas the FGLS and Harvey models always present the lowest AIC and BIC among the log-linear models. The Gamma always presents, for the three samples, the lowest AIC and BIC[16].

Finally, we do some cross-validation. We use the estimates of the gravity model for the year 1990 to calculate the predicted values in 1999. The fitted values are compared with the actual values in 1999. In order to compare log-linearized models with GLM, we use for the log-linear models the heteroskedasticity adjusted retransformation proposed by Baser (2007). Then the Root Mean Squared Predictor Error (RMSPE) is calculated. The results are summarized at the bottom of Tables 3 and 4 (first and second part) under RMSPE (1999). The RMSPE of different models can be compared directly, the model producing the smaller RMSPE being the better one. We repeat the same experiment for the three samples. Our findings indicate that for the first dataset (Table 3) the Heckman and Harvey models present the smaller RMSPE, for the second dataset (first part of Table 4, last row) the PPLM present the smaller RMSPE and for the third (second part of Table 4, last row) the NLS present the smaller RMSPE.

A matter of further research, outside the scope of the paper, is to check the accuracy and goodness of fit of two-part models. These models, the first part consisting of a probit or logit equation to distinguish between zero and positive outcomes and the second part using log-log models or GLM models on the positive values, allows different parameters and variables to determine the two parts of the data-generating process. Those models could be more appropriate than the Heckman model when the zeros represent outcomes of interest, as in the case of trade flows. This is recommendable when working with sectoral data, for which the

---

[16] Although AIC and BIC are not appropriate, strictly speaking, to evaluate models estimated by pseudo maximum likelihood, as stated by Santos Silva (personal communication), the criteria available based on the pseudo-likelihood function (PAIC, PBIC) add a penalty term that is very unstable and makes the PAIC and PBIC very inaccurate.

amount of zero values in export flows is much higher than when working with aggregated trade data.

5. Conclusions

The extended use of gravity models to predict international trade flows has generated an ongoing discussion concerning the estimation techniques to be applied. In particular, the issues concerning zero trade values and heteroskedastic residuals are controversial. Both issues also extend to gravity models used in social sciences to predict certain behaviour that mimic gravitational interactions, namely migration, or capital flows. Immigration between pairs of countries may be zero in a substantial percentage of observations, and omitting those zero observations biases the regression results. Bilateral flows of international migrants also exhibit tremendous variance both across destination countries and over time. Bilateral foreign direct investment data contains also many zero values and the residuals of the related models are in general heteroskedastic. This paper compares a number of estimation techniques that have been proposed in the economic literature to deal with both problems. The results of the simulations indicate that in these settings, PPML shows a similar performance to FGLS and Gamma estimates when the bias and the expected loss are used as the criteria to evaluate the performance of different estimators. The results also indicate that although the PPML is less affected by heteroskedasticity than other estimators, its performance is similar, in terms of bias and standard errors, to the FGLS estimator performance and in the case of no-zeros in the dependent variable, the GPLM present the lowest biases.

Nevertheless, it is worth noting that conclusions drawn from this type of simulation studies should be made with care, and that the winning estimation method should not be labelled "the new workhorse" for the estimation of constant-elasticity models. We have shown that small changes of the simulation setting can lead to different outcomes. Hence, one should be cautious and inspect each applied situation carefully in order to find the appropriate estimation method.

The results obtained when estimating the gravity model of trade using three different samples indicate that finding the best estimator for a given dataset requires carrying out a large number of tests and specification checks.

The use of transformed least squares methods trades robustness for some improvement in efficiency, whereas the use of GLM methods provides an easy way to deal with heteroskedastic errors but the researcher should choose the best variance function and in some cases the best GLM model cannot be estimated with adequate precision. These considerations are not only valid for gravity models of trade; they can also be extended to gravity models used to explain other types of inter-regional and international flows, including migration, development aid, investment and transport.

References

Anderson, J. E. and Wincoop, E.(2003) "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review* 93 (1), 70-192.

Baier, S. L. and Bergstrand, J. H. (2007) "Do Free Trade Agreements Actually Increase Members´ International Trade" Journal of International Economics 71, 72-95.

Baser, O. (2007) "Modeling Transformed Health Care Cost with Unknown Heteroskedasticity" Applied Economic Research Bulletin 01, 1-9.

Carroll, R.J. (1982) "Adapting for heteroskedasticity in linear models", *Annals of Statistics* 10, 1224-1233.

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

Chen, I-H. and Tsai, Y-Y.(2005) "Estimating the Staged Effects of Regional Economic Integration on Trade Volumes", National University of Kaohsiung, Taiwan. Mimeograph Document.

Deb, P., Manning, W. and Norton, E. (2008) "Modeling Health Care Costs and Counts" ASHE-Duke University Conference.

Delgado, M.A. (1992) "Computing Nonparametric Functional Estimates in Semiparametric Problems", *Econometric Reviews* 12(1), 125-128.

Delgado, M.A. (1993) "Semiparametric Generalized Least Squares in the Multivariate Nonlinear Regression Model", *Econometric Theory* 8, 203-222.

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

Greene, W.H. (2000) *"Econometric Analysis"*, Fourth Edition, Prentice Hall International, New Jersey.

Harvey, A. C. (1976) "Estimating regression models with multiplicative heteroscedasticity" *Econometrica* 44, 461-465.

Heckman, J.J. (1979) "Sample Selection Bias as a Specification Error", *Econometrica* 47, 153-161.

Helpman, E., M. Melitz and Y. Rubinstein (2008) "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics* 123 (2), 441-487.

Márquez-Ramos, L. Martínez-Zarzoso, I. and Suárez-Burguet, C. (2007) "The Role of Distance in Gravity Regressions: Is There Really a Missing Globalization Puzzle?, *The BE Journal of Economic Analysis and Policy* 7 (1). Article 6, 1-25.

Manning, W. G. (1998) "The Logged Dependent Variable, Heteroskedasticity and the Retransformation Problem" *Journal of Health Economics* 17, 283-295.

Manning, W. G. and Mullahy (2001) "Estimating Log Models: to Transform or not to Transform?", *Journal of Health Economics* 20, 461-494.

Martin, W. and Pham, C.S. (2008) "Estimating the Gravity Model when Zero Trade Flows are Frequent", Deakin University, Australia. Mimeograph document.

Martínez Zarzoso, I., F. Nowak-Lehmann D. and N. Horsewood (2009) "Are Regional Trading Agreements Beneficial? Static and Dynamic Panel Gravity Models", *North American Journal of Economics and Finance 20*, 46-65..

Martínez-Zarzoso, Inmaculada, Pérez-García, Eva and Suárez Burguet, Celestino (2008) "Do Transport Costs Have a Differential Effect on Trade at the Sectoral Level?" *Applied Economics* 40 (24), 3145-3157. Warwick, UK.

Oguledo, V.I. and Macphee, C.R. (1994), "Gravity Models: A Reformulation and an Application to Discriminatory Trade Arrangements" *Applied Economics* 26: 107-120. Warwick, UK.

Porojan, A. (2001) "Trading Flows and Spatial Effects: The Gravity Model Revisited", *Open Economies Review* 12, 265-280.

Robinson, P.M. (1987) "Asymptotically Efficient Estimation in the Presence of Hetereoskedasticity of Unknown Form", *Econometrica* 55(4), 875-891.

Rose, A.K. (2005) "Does the WTO Makes Trade More Stable?, *Open Economies Review* 16 (1), 7-22.

Sanso M., R. Cuairan, F. Sanz (1993) "Bilateral Trade Flows, the Gravity Equation, and Functional Form" *The Review of Economics and Statistics* 75 (2), 266-275.

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

Santos Silva, J.M.C. and Tenreyro, S. (2011) "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator", *Economics Letters*, forthcoming.

Soloaga, I. and Winters, A. (2001) "Regionalism in the Nineties: What Effects on Trade?", *North American Journal of Economics and Finance* 12 (1), 1-29.

Wooldridge, J.M. (2009). Introductory Econometrics, A Modern Approach, 4th ed., Cincinnati (OH): South-Western.

Table 1. Simulation results, regular case

	absolute error loss		bias		standard deviation	
	β1	β2	β1	β2	β1	β2
Case 1						
OLS (y>0)	0.3894	0.3560	0.3894	0.3560	0.0390	0.0529
NLS	**0.0059**	**0.0137**	**0.0000**	**0.0002**	**0.0075**	**0.0173**
Gamma	0.0535	0.0652	0.0139	0.0084	0.0672	0.0827
Poisson	0.0121	0.0215	-0.0001	0.0000	0.0161	0.0274
FGLS (y>0)	0.0943	0.0879	0.0795	0.0504	0.0737	0.0956
Case 2						
OLS (y>0)	0.2105	0.1992	0.2105	0.1992	0.0294	0.0497
NLS	0.0259	0.0452	0.0003	0.0012	0.0327	0.0569
Gamma	0.0336	0.0508	0.0043	0.0023	0.0431	0.0638
Poisson	**0.0156**	**0.0321**	**0.0001**	**0.0000**	**0.0196**	**0.0401**
FGLS (y>0)	0.0571	0.0931	0.0164	0.0008	0.0751	0.1285
Case 3						
OLS (y>0)	**0.0213**	**0.0424**	**0.0001**	**-0.0001**	**0.0267**	**0.0532**
NLS	0.2734	0.2331	0.1295	0.0255	3.4097	0.7309
Gamma	0.0253	0.0522	0.0003	-0.0006	0.0315	0.0650
Poisson	0.0551	0.0807	-0.0027	-0.0008	0.0721	0.1021
FGLS (y>0)	0.0219	0.0450	-0.0003	0.0001	0.0274	0.0564
Case 4						
OLS (y>0)	0.1327	0.1280	0.1326	-0.1247	0.0385	0.0754
NLS	0.9386	0.5972	0.7411	0.1663	20.5992	9.6913
Gamma	**0.0463**	**0.0855**	**0.0030**	**-0.0022**	**0.0589**	**0.1073**
Poisson	0.0777	0.1151	-0.0062	-0.0055	0.1030	0.1466
FGLS (y>0)	**0.0591**	**0.1017**	0.0492	-0.0881	**0.0506**	**0.0861**

Note: The absolute error loss is defined in Equations 7 and 8 in the main text.

Table 2. Simulation results with zeros in the dependent variable

	absolute error loss		bias		standard deviation	
zeros at random	β1	β2	β1	β2	β1	β2
Case 1						
OLS (y>0)	0.3900	0.3556	0.3900	0.3556	0.0429	0.0586
NLS	0.0694	0.0889	0.0043	0.0080	0.0907	0.1226
Gamma	0.0583	0.0721	0.0154	0.0072	0.0738	0.0920
Poisson	**0.0290**	**0.0422**	**-0.0003**	**-0.0003**	**0.0379**	**0.0530**
FGLS (y>0)	0.0979	0.0903	0.0826	0.0539	0.0760	0.0975
Case 2						
OLS (y>0)	0.2104	0.1993	0.2104	0.1993	**0.0322**	**0.0529**
NLS	0.0743	0.1038	0.0045	0.0073	0.0963	0.1371
Gamma	0.0372	0.0570	0.0045	0.0026	0.0473	0.0719
Poisson	**0.0299**	**0.0486**	**-0.0011**	**0.0004**	0.0383	0.0610
FGLS (y>0)	0.0575	0.0941	0.0195	0.0033	0.0749	0.1301
Case 3						
OLS (y>0)	0.0225	0.0462	-0.0003	0.0007	0.0282	0.0577
NLS	0.4154	0.3285	0.2529	0.0829	5.6568	2.4460
Gamma	0.0292	0.0596	-0.0002	0.0004	0.0366	0.0750
Poisson	0.0643	0.0937	-0.0036	-0.0028	0.0836	0.1185
FGLS (y>0)	**0.0237**	**0.0483**	**0.0001**	**0.0008**	**0.0299**	**0.0607**
Case 4						
OLS (y>0)	0.1322	0.1299	0.1321	-0.1254	0.0417	0.0830
NLS	0.8055	0.5406	0.5860	0.1639	7.8462	3.2434
Gamma	**0.0495**	**0.0976**	**0.0046**	**0.0000**	0.0626	0.1229
Poisson	0.0863	0.1281	-0.0084	-0.0045	0.1149	0.1649
FGLS (y>0)	**0.0613**	**0.1053**	0.0500	-0.0878	**0.0536**	**0.0941**
zeros with pattern	β1	β2	β1	β2	β1	β2
Case 1						
OLS (y>0)	0.3408	0.2877	0.3408	0.2877	0.0448	0.0520
NLS	**0.0116**	**0.0140**	**0.0112**	**0.0009**	**0.0082**	**0.0176**
Gamma	0.2651	0.0801	0.2620	0.0097	0.0967	0.1014
Poisson	0.0669	0.0221	0.0668	0.0008	0.0166	0.0279
FGLS (y>0)	0.0760	0.0765	0.0571	0.0416	0.0714	0.0862
Case 2						
OLS (y>0)	0.1938	0.1776	0.1938	0.1776	0.0336	0.0488
NLS	**0.0273**	**0.0468**	0.0118	0.0021	0.0330	0.0586
Gamma	0.2490	0.0634	0.2487	0.0027	0.0644	0.0799
Poisson	0.0664	0.0335	0.0664	0.0010	**0.0205**	**0.0420**
FGLS (y>0)	0.0542	0.0889	**0.0111**	**0.0054**	0.0730	0.1227
Case 3						
OLS (y>0)	**0.0248**	**0.0465**	**0.0002**	**-0.0007**	**0.0311**	**0.0583**
NLS	0.3080	0.2827	0.1818	0.0921	3.6995	2.0455
Gamma	0.2405	0.0622	0.2405	0.0011	0.0408	0.0777
Poisson	0.0736	0.0834	0.0633	-0.0027	0.0716	0.1054
FGLS (y>0)	**0.0258**	**0.0491**	0.0001	0.0010	0.0325	0.0616
Case 4						
OLS (y>0)	0.1200	0.1545	0.1199	-0.1529	0.0448	0.0794
NLS	0.5178	0.4039	0.3373	0.1053	4.9428	3.1924
Gamma	0.2512	0.0999	0.2508	-0.0011	0.0773	0.1261
Poisson	0.0870	0.1194	0.0601	-0.0081	0.1047	0.1521
FGLS (y>0)	**0.0584**	**0.1078**	0.0465	0.0929	**0.0540**	**0.0914**

Note: The absolute error loss is defined in Equations 7 and 8 in the main text.

Table 3. Estimation results for 1990: traditional gravity model

	LOG-LOG MODELS			TWO PART HECKMAN		GENERALISED LINEAR MODELS		
	OLS	FGLS	Harvey	Heckman		Poisson	Gamma	NLS
Dep. Var.	log(exports)	log(exports)	log(exports)	log(exports)	exports	exports	exports	exports
lyi	1.08	0.932	1.077	1.069	0.38	0.799	0.739	0.86
	0.016	0.009	0.014	0.017	0.012	0.037	0.037	0.07
lyj	0.951	0.899	0.958	0.948	0.144	0.798	0.744	0.91
	0.016	0.008	0.014	0.016	0.011	0.03	0.032	0.079
lyhi	0.777	0.791	0.793	0.774	0.143	0.549	0.596	0.587
	0.028	0.019	0.025	0.027	0.019	0.08	0.043	0.129
lyhj	0.585	0.574	0.552	0.58	0.218	0.614	0.462	0.745
	0.029	0.016	0.025	0.027	0.019	0.054	0.055	0.183
ldist	-1.326	-0.903	-1.226	-1.316	-0.398	-0.713	-1.066	-0.797
	0.033	0.015	0.029	0.033	0.026	0.038	0.043	0.119
border	0.714	0.26	0.554	0.712	0.398	0.427	0.457	0.194
	0.136	0.04	0.132	0.154	0.162	0.113	0.151	0.124
comlang	0.329	0.335	0.396	0.325	0.085	0.585	0.211	0.551
	0.062	0.029	0.054	0.06	0.043	0.109	0.078	0.333
colony	1.635	0.679	1.377	1.627	8.598	0.101	1.219	-0.11
	0.112	0.042	0.124	0.165	8.14E+06	0.144	0.125	0.428
landl	-0.187	-0.3	-0.206	-0.19	0.059	-0.589	-0.268	-0.091
	0.051	0.027	0.046	0.05	0.034	0.108	0.092	0.474
island	0.008	0.354	0.048		0.126	0.391	-0.016	0.652
	0.05	0.024	0.046		0.035	0.077	0.069	0.18
landap	-0.047	-0.052	-0.058	-0.047	-0.027	-0.057	-0.016	-0.001
	0.011	0.005	0.009	0.01	0.007	0.024	0.019	0.066
regional	0.915	0.15	0.443	0.918	0.517	0.077	0.459	0.36
	0.145	0.037	0.155	0.171	0.187	0.093	0.151	0.156
custrict	1.714	0.698	1.851	1.691	0.529	-0.084	0.689	-0.015
	0.185	0.047	0.204	0.236	0.195	0.203	0.207	0.433
comcol	0.664	0.766	0.624	0.662	0.17	0.811	0.643	1.671
	0.101	0.076	0.085	0.085	0.053	0.296	0.121	1.124
_cons	-37.875	-36.377	-38.289	-37.5	-11.26	-28.51	-22.46	-36.549
	0.685	0.342	0.537	0.628	0.448	1.231	0.76	5.546
R-squared	0.619	0.862	0.679					
N	12134	12134	12134	13974		13974	13974	13974
ll	-28189.5	-18299.6	-26833.8	-32312.3		-7.90E+11	-234853	-302957
rmse	2.4716	1.093957	2.210317	2.5114				
aic	56409.08	36629.1	53697.61	64686.49		1.13E+08	33.615	43.36218
bic	56520.13	36740.15	53808.67	64920.39		1.58E+12	-86841.5	5.54E+21
RMSE retrans	1.81E+09	1.87E+09	1.70E+09	1.56E+09		7.39E+08	2.31E+09	6.30E+08
RMSEP	9.94E+10	3.70E+09	3.40E+09	3.26E+09		4.22E+09	1.26E+10	5.61E+09

Note: Ly_i (Ly_j) and Lyh_i (Lyh_j) denote income and income per capita in country i (j), respectively. Border is a binary variable that is unity if countries *i* and *j* share a common border, 0 otherwise. *Comlang* is a binary variable that is unity if countries i and *j* share a common language, 0 otherwise. *Colony* is a dummy that takes the value of 1 if country *i* ever colonized *j* or *vice versa*, 0 otherwise. Island takes the value of 1 if country *i or country j or both* are islands. *Landl* is the number of landlocked countries in the country pair. *Landap* is the total area of both countries, *i* and *j*. *FTA* is a dummy that takes the value of 1 when the trading partners belong to the same trade agreement, 0 otherwise. *Custrict* is a dummy that takes the value of 1 when the trading partners share a common currency, 0 otherwise. *Comcol* is a dummy that takes the value of 1 if *i* and *j* were ever colonies after 1945 with the same colonizer, 0 otherwise. Standard errors reported below each estimate.

Table 4. Estimation results for 1990: Theoretically justified gravity model

65 Countries'	LOG-LINEAR MODELS			TWO PART HECKMAN		GENERALIZED LINEAR MODELS		
	OLS	FGLS	Harvey	Heckman		Poisson	Gamma	NLS
Dep. Var.	log(exports)	log(exports)	log(exports)	log(exports)	exports	exports	exports	exports
Ldist	-1.134	-1.121	-1.081	-1.095	-0.97	-0.539	-1.191	-0.454
	0.045	0.042	0.04	0.039	0.078	0.044	0.057	0.085
Comlang	0.639	0.587	0.645	0.623	0.592	0.271	0.946	0.448
	0.085	0.077	0.073	0.084	0.136	0.064	0.101	0.101
Border	0.393	0.113	0.245	0.402		0.658	0.533	0.588
	0.151	0.136	0.137	0.148		0.089	0.168	0.187
FTA	-0.064	-0.106	-0.073		0.172	0.287	-0.017	0.540
	0.045	0.041	0.037		0.120	0.050	0.056	0.072
Constant	12.824	15.036	14.236	12.388	6.429	11.761	16.464	10.335
	0.55	0.573	0.542	0.505	0.771	0.559	0.693	0.863
R-squared	0.801	0.835	0.853		Deviance	4.96E+8	5168	2.01E+15
RMSE	1.395448	1.182354	1.14222		Dispersion	5.92e+8	13158	2.01E+15
Nobs	3230	2938	3230	3220	3804	3804	3804	3804
Ll	-5593.633	-4595.019	-4946.855	-6211.515		-2.48E+08	-41675.05	-56731.24
AIC	11445.27	9448.038	10151.71	12939.03		130566	21.979	29.89497
BIC	12229.62	10220.17	10936.06	14549.93		4.97E+08	-25127.24	2.00E+15
RMSE retrans	12056872	4673505.5	3431688	8477757		978886.81	22653696	725514.25
RMSPE (1999)	15804256	8104227.5	8640781	7839773.5		3661013.5	20849120	4846363
47 countries'	OLS	FGLS	Harvey	Heckman		Poisson	Gamma	NLS
Dep. Var.	log(exports)	log(exports)	log(exports)	log(exports)	exports	exports	exports	exports
Ldist	-1.064	-1.04	-1.06	-0.841	-0.705	-0.564	-1.396	-0.294
	0.058	0.044	0.05	0.09	0.055	0.052	0.069	0.078
Island	-0.37	-0.284	-0.279	-0.099	-0.161	-1.034	-1.024	-0.825
	1.529	0.877	0.435	1.612	0.076	0.403	0.472	0.469
Comlang	1.416	1.184	1.259	1.23	0.51	0.775	1.74	0.745
	0.122	0.089	0.103	0.141	0.119	0.069	0.147	0.07
Border	-0.354	-0.347	-0.349		-1.526	0.246	-0.218	0.213
	0.215	0.162	0.176		0.258	0.079	0.207	0.077
EU	-0.369	-0.462	-0.461	-0.479	5.722	0.04	-1.729	-0.189
	0.191	0.122	0.142	0.199	.	0.09	0.199	0.101
NAFTA	1.071	0.478	0.657	0.668	6.476	1.322	-0.037	2.571
	0.642	0.426	0.598	0.671	.	0.184	0.456	0.363
CACM	1.805	1.765	1.841	1.884	-0.118	1.481	1.245	1.909
	0.388	0.246	0.281	0.401	0.559	0.34	0.328	0.84
CARIC	1.393	1.44	1.447	1.763	-1.205	0.917	2.158	0.662
	0.231	0.233	0.228	0.255	0.13	0.331	0.289	0.9
Constant	11.968	15.402	12.092	10.185	6.797	18.786	26.599	17.334
	1.677	1.426	0.822	0.801	0.485	0.457	0.654	0.587
R-squared	0.817	0.903	0.878		Deviance	1.78E+8	3153	5.67E+14
RMSE	1.468021	1.005774	0.0003035		Dispersion	1.99E+8	7910	5.67E+14
Nobs	1656	1656	1656	2162		2162	2162	2162
Ll	-2933.42	-2307.191	11116.21			-8.91E+07	-21359.71	-31491.06
AIC	6068.839	4816.382	-22032.42	.		8.24E+04	19.85265	29.22484
BIC	6615.467	5363.011	-21491.21	.		1.78E+08	-12672.58	5.68E+14
RMSE retrans	13452503	6886919.5	9055090	12829089		666638.63	61030884	512352.28
RMSPE (1999)	11174116	5939834	7465615.5	10453723		4320483	58862776	4307895.5

Note: The effects of income and income per capita variables cannot be estimated since exporter and importer effects are added as regressors. Border, Comlang, Island and FTA are defined as in Table 3. EU, NAFTA, CACM and CARICOM denote trade creation effects. Standard errors reported below each estimate.

Table 5. Estimation results for panel data (1980-1999)

47 countries' Sample	LOG-LINEAR MODELS			GENERALIZED LINEAR MODELS		
	OLS	FGLS, het	FGLS, het ar1	Poisson	Gamma	NLS
Dep. Var.	log(exports)	log(exports)	log(exports)	exports	exports	exports
Lyi	0.941	0.889	0.907	0.782	0.848	0.759
	0.006	0.012	0.014	0.036	0.040	0.043
Lyj	0.889	0.849	0.824	0.824	0.698	0.701
	0.005	0.01	0.01	0.034	0.034	0.039
Lyhi	0.870	0.435	0.915	0.191	1.025	-0.033
	0.036	0.065	0.073	0.185	0.152	0.111
Lyhj	0.217	-0.367	-0.300	-0.205	0.390	-0.256
	0.012	0.017	0.016	0.052	0.04	0.069
Ldist	-1.176	-1.276	-1.533	-0.618	-0.941	-0.470
	0.012	0.021	0.024	0.067	0.065	0.04
Border	-0.061	0.006	0.091	0.448	0.348	0.478
	0.028	0.049	0.071	0.115	0.211	0.108
Comlang	0.896	0.721	0.629	0.625	1.185	0.618
	0.023	0.049	0.057	0.113	0.279	0.086
Island	-0.048	-0.021	0.012	0.055	-0.075	0.12
	0.019	0.039	0.047	0.115	0.198	0.111
EU	0.952	0.521	0.561	0.339	0.889	-0.081
	0.026	0.031	0.032	0.123	0.151	0.095
EUX	0.125	-0.270	-0.160	-0.522	0.113	-0.510
	0.022	0.031	0.028	0.163	0.121	0.207
EUM	0.902	0.864	0.998	0.252	0.827	-0.052
	0.033	0.052	0.062	0.179	0.184	0.193
NAFTA	0.702	0.171	0.037	0.468	1.437	0.650
	0.182	0.127	0.123	0.257	0.354	0.153
NAFTAX	-0.288	-0.137	-0.140	-0.551	-0.397	-0.621
	0.03	0.031	0.03	0.132	0.134	0.113
NAFTAM	0.463	0.326	0.425	0.163	0.384	-0.206
	0.04	0.057	0.061	0.256	0.132	0.178
Constant	-35.748	-23.336	-26.421	-23.470	-32.502	-17.947
	0.37	0.663	0.721	2.05	1.851	2.154
Nobs	12507	12506	12506	13264	13264	13264
Deviance				6.10E+09	24015.77	4.62E+16
Chi2_dev				7.18E+09	50382.26	4.62E+16
Dispersion					3.798421	3.48E+12
Chi2_dis				541439.1	3.798421	3.48E+12

Note: Ly_i (Ly_j) and Lyh_i (Lyh_j) denote income and income per capita in country i (j), respectively. Ldist, Border, Comlang and Island are defined as in Table 3. Ldist, Border, Comlang and Island are defined as in Table 3. EU and NAFTA denote trade creation effects, EUM, NAFTAM, CACMM and CARICOMM denote import-diversion effects; and EUX and NAFTAX denote export diversion effects. Standard errors reported below each estimate.

Table 6. Test for the implicit assumption in PPML and OLS/FGLS (Rose data, various years)

PARK TEST LOG-LOG	1980	1990	1995	2000
λ_1	1.803	1.829	1.834	1.856
95% Confidence Interval	1.786,1.820	1.816,1.842	1.822,1.846	1.846,1.865
St. error	0.011	0.008	0.007	0.007
Nobs	11852	13974	17616	15110
PARK TEST GAMMA	1980	1990	1995	2000
λ_1	1.68	1.586	1.639	1.646
95% Confidence Interval	1.534,1.826	1.451,1.722	1.571,1.708	1.472,1.819
St. error	0.074	0.069	0.035	0.088
Nobs	11852	13974	17616	15110
(Rose Data)				

Note: λ_1 determines the relationship between raw-scale mean and variance functions.

Appendix

A.1 Simulations 40 percent zeros

40% Zeros		Absolute error loss		Bias		Standard deviation	
		β1	β2	β1	β2	β1	β2
Case 1	**Poisson**	0.0673	0.0221	0.0673	0.0010	0.0002	0.0003
	FGLS	0.0944	0.0873	0.0787	0.0507	0.0008	0.0009
Case 2	**Poisson**	0.0539	0.0788	-0.0018	-0.0014	0.0007	0.0010
	FGLS	0.0621	0.0989	0.0231	0.0054	0.0008	0.0014
Case 3	**Poisson**	0.0850	0.1237	-0.0068	-0.0028	0.0011	0.0016
	FGLS	0.0289	0.0574	-0.0003	0.0000	0.0004	0.0007
Case 4	**Poisson**	0.1066	0.1605	-0.0114	-0.0101	0.0014	0.0020
	FGLS	0.0667	0.1159	0.0504	-0.0884	0.0006	0.0011