Are estimation techniques neutral to estimate gravity equations? An application to the impact of EMU on third countries' exports.

Abstract

The gravity equation has been traditionally used to study the determinants of trade flows across countries. However, several problems related with its empirical application still remain unclear. In this paper, we provide a survey of the literature concerning the specification and estimation's method of this equation in last years. Additionally, we test the fit of different estimation procedures (Poisson, panel) using a large database. Our second objective is to assess the effect of the EMU on non EU countries exports, a question that hasn't been clearly answered until now.

* Classification JEL Code. C13, C33, F10, F15

Authors

Estrella Gómez Herrera (Universidad de Granada)

Juliette Milgram Baleix (Universidad de Granada)

Dto de Teoría e Historia económica

Facultad de Ciencias Económicas y Empresariales.

Universidad de Granada.

Campus de Cartuja, s/n.

18.011. Granada (Spain).

Telf.: +34 958 24 10 00 Ext 20145

Fax: +34 958 24 99 95.

Estrella Gómez [estrellagh@gmail.com]

Juliette Milgram Baleix [jmilgram@ugr.es]

1. Introduction

In the last fifty years, the gravity equation of trade has been widely used to predict trade flows. After the controversies concerning its theoretical foundation in the eighties and about its specification in the nineties, the estimation of gravity models went through an intense debate about estimations techniques in last years. Traditionally the multiplicative gravity model was linearised and estimated using OLS techniques, assuming that the variance of the error is constant across observations (homoscedasticity) or using panel techniques, assuming that the error is constant across countries or country-pairs. As pointed by Silva and Tenreyro (2006) in presence of heteroscedasticity, the Pseudo Poisson Maximum Likelihood (PPML) estimator performs better since OLS is not efficient. Another challenge of this literature concerns the zero values. Helpman et al. (2008) renewed this debate by proposing a theoretical foundation of these zero values based on a model with heterogeneity of firms à la Melitz 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), Siliverstov and Schumacher (2007), Westerlund and Wilhelmsson (2007) have obtained some divergent results when comparing alternative estimators to deal with the heteroscedasticity and zero values problems.

The aim of this paper is twofold. Our prior objective is to contribute to the methodological debate on heteroscedasticity in three dimension (i, j, t) datasets and compare several estimation techniques. To this end, we use a gravity equation based on Anderson and van Wincoop (2003)'s model. We discuss the fit of different estimation procedures applied to a large dataset of bilateral exports for 47 countries (80% of world trade) over the period 1980-2002.

Our second objective is to assess the effect of EMU on non EU countries exports, a question that hasn't been clearly answered until now. The sensitiveness of exports to exchange-rate regimes - defined in a de facto way by the level and the volatility of the exchange rate – is also explored. Additionally, we test how the euro affects trade among EMU countries and its imports from third countries by introducing dummies reflecting the fact that one, or both partners, belongs to the EMU. We compare how the coefficients of volatility of exchange rates and the dummies for trade and monetary agreements are affected by the different estimation techniques.

There is little debate about trade flows being determined by the behaviour of real exchange rates: even when market structures are taken into account (for instance when they give rise to pricing to market strategies) an appreciation in the real exchange rate leads to a worsening of the competitive position of the economy, and consequently to a rise in imports, and a fall in exports. This fact is now well documented, and is robust to the use of alternative measurement strategies even if aggregate demand and supply elasticities also depend on the structure of specialization in each country. The impact of exchange rate volatility on trade is more controversial, both in theory and empirical analysis. In theory, an increase in exchange rate volatility could either increase or decrease trade, depending on the risk aversion of firms or on the shape of the production functions. Looking at empirical analysis suggests that the measured effects of exchange-rate volatility on trade can be either very low and little significant or significantly negative, though minor in magnitude. Though, monetary agreements may have an additional positive impact on trade flows once volatility reduction and exchange rate are controlled for as showed by Gil et al. (2009). Though, the question of the appropriate exchange rate strategy for the neighbors of the Eurozone and the impact it could have on third countries exports to these members is not completely solved.

To anticipate our most important findings, our study confirms that the estimation technique is not neutral to study the effect of exchange-rate regimes – defined by the level of and the volatility of the (real) exchange rate – on exports; though it doesn't matter so much for a basic model of trade flows. Different techniques lead to divergent results when the impact of EMU is studied. All in all, our results do not show strong diversion effects of the EMU.

The rest of the paper is organized as follows. In the next section we present the theoretical model. Section 3 details some of the most usual estimation methods in the gravity literature. In Section 4 the baseline model and the data are presented. Section 5 compares a new model with the baseline in order to assess the impact of exchange rate variables and EMU on trade. Some conclusions are provided in Section 6. The Figures and Tables are confined to the Appendix.

2. From the theory to the specification of the gravity equation

a) The model

The gravity equation of trade is highly effective at explaining bilateral flows as proven at a very early date by the works of Linnemann (1966) and Leamer and Stern (1971). However, this model threw several controversies. Theoretical framework was putted into doubt and afterwards justified: Bergstrand, 1989 for the factorial model; Deardorff, 1998 for the Hecksher-Ohlin model; Anderson, 1979 for goods differentiated according to their origin, and Helpman, Melitz and Rubinstein, 2008 in the context of heterogeneity of firms. It seems that the H-O model would better explain the success of the gravity equation when the partners have very different factorial endowments, while increasing returns models would better explain the exchanges between similar countries precisely because the exchanges of differentiated goods represent a significant share of their trade.

In this paper we consider the augmented version of the Anderson (1979) model proposed by Anderson and van Wincoop (2003). They assume that goods are differentiated by origin; that each country is specialized in the production of only one good and preferences are identical, homothetic and approximated by a constant elasticity of substitution (CES) function. A world where goods are differentiated by origin may fit well with a sample of countries that are not completely similar regarding endowments and demand but not too heterogeneous; so taste for varieties may play an important role. This model is overall interesting to the extent that the discussion of the multilateral resistance may matter for the heteroscedasticity considerations. The relationship goes as follows: if a country is remote, it will tend to diversify its production (since trade costs are higher).

However, if it is located near to other countries, specialization would be higher, and trade flows in that case would be more frequent. The variance of trade is then associated with one of the regressors; making heteroscedasticity to appear. As it is well-known, Anderson and van Wincoop (2003) argued that *"remoteness" variable related to distance to all bilateral partners was a key variable for gravity models.*

The utility function is stated as:

$$U_j = \left(\sum_i \beta_i^{(1-\sigma)/\sigma} c_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)} \tag{1}$$

where c_{ij} is the consumption of importer j in goods from exporter i and σ is the elasticity of substitution between goods from different countries and β is a positive parameter. Consumer's constraint is given by:

$$y_j = \sum_i x_{ij} = \sum_i p_{ij} c_{ij} \tag{2}$$

where y_j is the nominal income of importer j, p_{ij} is the c.i.f. import price of exporter i's goods for importer j consumers and x_{ij} is the nominal value of exports from i to j. Prices differ among countries due to trade costs that are not directly observable. Trade costs are modeled as Krugman's iceberg costs, which implies that t_{ij} units of good from export country i need to be shipped in order for one unit to reach country j. Given the exporter's supply price, p_i, the export price would be $p_{ij} = p_i t_{ij}$. This means that for each unit of good shipped from i to j the trade cost would be $t_{ij} - 1$ (in terms of the good i lost to shipping). The exporter passes on these trade costs to the importer. It should be noted that if $i = j$, then $t_{ij} = 1$ and $p_{ij} = p_i$.

Solving the maximization problem of consumer j, the nominal demand of country j for goods from country i is obtained as:

$$x_{ij} = \left(\frac{\beta_i p_i t_{ij}}{P_j} \right)^{(1-\sigma)} y_j \tag{3}$$

P_j is a function of j's full set of bilateral trade resistance terms:

$$P_j = \left[\sum_i \left(\beta_i p_i t_{ij} \right)^{1-\sigma} \right]^{1/(1-\sigma)} \tag{4}$$

where β is a positive distribution parameter that denotes the share of importer j in country i's consumption. This is the key innovation introduced by Anderson and van Wincoop (2003): they claim that bilateral trade costs is not the only factor affecting trade, but also trade costs of each country with all others. Hence, three components of trade resistance can be identified: bilateral trade barriers between region i and j, (t_{ij}); i's resistance to trade with other countries (P_i) and j's resistance to trade with other countries (P_j).

Imposing market clearing condition,

$$y_i = \sum_j x_{ij} = \sum_j (\beta_i t_{ij} p_i / P_j)^{(1-\sigma)} y_j = (\beta_i p_i)^{1-\sigma} \sum_j (t_{ij} / P_j)^{(1-\sigma)} y_j \tag{5}$$

Scaled prices, $\{\beta_i p_i\}$, are solved from this equation and substituted into (3). Defining $y^w \equiv \sum_j y_j$ as the world income and income shares by $\theta_j \equiv y_j / y^W$, equation (3) can be solved as:

$$x_{ij} = \frac{y_i y_j}{y^w} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma} \tag{6}$$

where

$$\Pi_i \equiv (\sum_j (t_{ij} / P_j)^{1-\sigma} \theta_j)^{1/(1-\sigma)} \tag{7}$$

Once $\{\beta_i p_i\}$ are substituted into (4), P_j can be expressed as:

$$P_j = (\sum_i (t_{ij} / \Pi_i)^{1-\sigma} \theta_i)^{1/(1-\sigma)} \tag{8}$$

Now, (7) and (8) can be solved for all Π_i's and P_i's in terms of income shares, θ_j, bilateral trade barriers, t_{ij}, and σ. If trade costs are assumed to be symmetric ($t_{ij} = t_{ji}$), solution to (7) and (8) can be demonstrated to be $\Pi_i = P_i$, with

$$P_j^{1-\sigma} = \sum_i P_i^{\sigma-1} \theta_i t_{ij}^{1-\sigma} \tag{9}$$

Hence, an implicit solution to the price indices can be obtained as a function of all bilateral trade barriers and income shares. The gravity equation becomes:

$$x_{ij} = \frac{y_i y_j}{y^w} \left(\frac{t_{ij}}{P_i P_j} \right)^{1-\sigma} \tag{10}$$

As it can be appreciated, what matters in this specification is the bilateral trade cost *relative to* an overall index of trade costs, that is, bilateral trade resistance compared to multilateral trade resistance. Taking into account the relative prices also implies that trade barriers reduce trade between (and within) large countries more than between (and within) small ones.

Since t_{ij} is not observed, is defined as a loglinear function of observable variables: bilateral distance and a dummy variable, b_{ij} that takes value 1 if *i* and *j* are located in different countries, and zero otherwise. Then:

$$t_{ij} = b_{ij} d_{ij}^{\rho} \tag{11}$$

Substituting this term in the initial equation and taking logarithms with an error term, ε_{ij}, and with a constant term, *k*, we would have a linear standard gravity equation:

$$\ln T_{ij} = k + \ln y_i + \ln y_j + (1-\sigma)\rho \ln d_{ij} + (1-\sigma)\ln b_{ij} - (1-\sigma)\ln P_i - (1-\sigma)\ln P_j + \varepsilon_{ij} \qquad (12)$$

b) Multilateral Trade Resistance (MTR)[1]

Anderson and van Wincoop's proposal was inspired by a pioneer article of McCallum (McCallum, 1995) who studies the importance of border effects. Actually, they use the same database. The claim of Anderson and van Wincoop was that McCallum (1995)'s equation suffered from omission of variables that might translate in an overestimation of border effects since he did not include a measure of the multilateral trade resistance. To solve that problem, they develop a new theoretical framework for the gravity equation that includes a theoretical specification for the multilateral resistance term. The specification is given by equation (8). Since the multilateral price indexes (P_i and P_j) are not observed, some alternatives have been proposed for estimation purposes.

A first option is to include price index data directly (Ruiz and Vilarubia 2007). This solution has never been used due to the lack of data or non homogeneity of the calculation methods among national sources.

A second solution, proposed by Anderson and vanWincoop (2003), is a non-linear estimation technique. To obtain the multilateral trade resistance terms, they use the observables in their model, (distances, borders, and income shares). Assuming symmetric trade costs, using 41 goods market-equilibrium conditions[2] and a trade cost function defined in terms of observables, they are able to obtain the P_i and P_j terms. They argue that this method is more efficient than any other. However, the procedure is data consuming and has not been frequently used by other authors.

A method frequently used is to include a proxy for these indexes called "remoteness variable":

$$Rem_i = \sum_j \frac{dist_{ij}}{(GDP_j/GDP_{ROW})}$$

where the numerator would be the bilateral distance among two countries, and the denominator would be the share between each country's GDP in the rest of the world's GDP. Anderson and vanWincoop (2003) compare their previous results with a regression including his remoteness variable. They claim that this procedure is not theoretically correct, since the only trade barrier the variable captures is distance. Even if distance were actually the only bilateral barrier, they argue that the way in which it is included in the remoteness index is not theoretically justified.

Head and Mayer's (2000) remoteness variable describes the full range of potential

[1] The most frequent empirical approach to measure the border effect is the gravity equation. Other studies on border effects are Wei (1996), Helliwell (1998), Evans (2000), Head and Mayer (2000), Feenstra (2002), Gil-Pareja et al. (2005) or Cafiso (2008).

[2] In their sample, they use the same 30 US states and 10 Canadian provinces that McCallum (1995) includes. There are 20 additional states, plus Columbia, that they aggregate into one. Finally, they have 41 equations.

suppliers to a given importer, taking into account their size, distance and relevant costs of crossing the border[3].

The method most commonly used is the one proposed by Feenstra (2002). It consists in including importer and exporter fixed effects in order to control for the specific country multilateral resistance term, instead of estimating it. The coefficient of the dummies for the importer and the exporter should reflect the multilateral resistance of each country. We will detail this procedure in section 3. In table A1 we detail several studies that implemented this technique, complemented by the time dimension; for instance Micco, Stein and Ordoñez (2003); Baltagi, Egger and Pfaffermayr (2003); Cheng and Wall (2005); Glick and Rose (2001); Ruiz and Vilarrubia (2008); Vicarelli and Benedictis (2004); Fidrmuc (2008); or Henderson and Millimet (2008).

Finally, Baier and Bergstrand (2006) suggest generating a linear approximation of the P_i and P_j terms by means of a first-order Taylor series expansion. This procedure is a little more complicated than simply including fixed effects, but it avoids the non-linear procedure employed by Anderson and vanWincoop (2003), and allows for estimation within OLS. Baier and Bergstrand's method is theoretically consistent and captures country specific and country-pair specific effects.

c) Augmenting gravity equation

Concerning the proxy for supply and demand sizes used in the gravity equation, the most common feature is to use GDP for the importer and for the exporter. In some cases GDP per capita is also introduced as a proxy for capital-labour intensities (not only factor endowments of a country).

Additionally, it is commonly accepted that geographical distance may be a poor approximation of all the economic barriers for international trade. To control better these omitted variables, the general gravity equation proposed above has been completed by a wide range of variables depending on the focus of the paper. It is common to include:

- Adjacency. This variable takes value 1 if trade partners share a common border. The effect of this variable on trade is expected to be positive.

- Common language: sharing a language should make all transaction easier and costless.

- Colonial links: this effect is introduced by means of a dummy variable. There are two different aspects that may be included: to have had a common colonizer or to have been colonized by the other country in the past. In both cases, the expected influence is positive since a colonial relationship is prone to reduce cultural differences and costumes between two countries.

- Religion: this variable takes value 1 if both countries share the same religion. It is expected to have a positive effect over trade.

- RTA: This effect has been widely studied due to the proliferation of these agreements in the last 20 years. Some articles related with this issue are Frankel, Stein and

[3] Wei (1996), Wolf (1997), and Helliwell (1997) are other examples of regressions that include a remoteness variable.

Wei (1995), Sapir (2001), Soloaga and Winters (2001), Greenaway and Milner (2002), Martínez-Zarzoso et al. (2003), Fratianni and Hoon Oh (2007) or Hoon Oh and Travis Selmier (2008). These authors try to answer the question of the existence of a regional bias in trade, and to discern the trade potential associated with integration. The effect of an RTA over trade is captured by the introduction of dummies. Other specifications (Baldwin et al. 2006) introduce different dummies for the cases in which none, both or only one of the two countries belong to the agreement.

We could mention some other variables that are not so frequently included:

- Technological variables: the influence of these variables on trade is increasing through time. Márquez-Ramos et al. (2005) include them in their model using a different specification for the "hard" and the "soft" investment in infrastructure in a country. Freund and Weinhold (2004) also include technological variables in their specification.

- Access to water: The relationship is again positive, since access to water reduces transport costs. Some specifications of the gravity equation include a "landlocked" or “island” variable to capture a similar effect.

- Area: The bigger the country, the lower the necessity it would have of importing, so the effect of this variable on trade is negative.

- Exchange rate: this variable is part of the determinants of trade volume and it is expected to affect trade. Curiously, it is not frequently incorporated in the gravity equation.

-Exchange rate volatility. Frankel and Wei (1995, 1997) evidence a significant negative impact of exchange-rate volatility on trade flows across Asian countries on a cross-section basis, a result found to be strongly robust by Rose (2000), who finds exchange-rate volatility to be a significant and systematic impediment to trade for an extensive sample of countries. Gil-Pareja et al. (2008) analyze the impact of this variable in a sample of 25 countries over the period 1950-2004 and finds also a statistically significant negative effect on trade. Finally, Tenreyro (2006) finds opposite results. Following Santos Silva and Tenreyro (2006), she uses pseudo-maximum likelihood (PML) technique to deal with heteroskedastic biases. To deal with the endogeneity and the measurement error of exchange rate variability she then develops an instrumental-variable (IV) version of the PML estimator. Results indicate that nominal exchange rate variability has no significant impact on trade flows.

3. Estimation methods

The new workhorse in the estimation of the gravity equation is still unclear. Every method presents important advantages and disadvantages: some of them solve the heteroscedasticity or the zero problems but are too costly, whereas other simpler methods are not useful in the presence of those two characteristics, or do not take into account the multilateral dimension of trade.

For that reason, becomes a frequent practice in the literature to include several estimation methods using the same database, in order to check which one performs better. We will describe the most important ones, and include them in our estimation.[4] It is also frequent to check this performance with Monte Carlo simulations (Silva and Tenreyro 2006, Martínez-Zarzoso et al. 2007, Martin and Pham 2008, etc.). We detail some of the results in the appendix (Table A2).

Recently, the problem of the zero flows has been revisited. The literature distinguishes several methods of dealing with that problem. Truncation (elimination) or censoring methods have been widely used. However, these methods have not a strong theoretical support and do not guarantee consistent estimates, so they have not been employed frequently in the literature. Alternative solutions are Tobit estimation, Poisson Pseudo Maximum Likelihood estimation, Nonlinear Least Squares (NLS), Feasible General Least Squares (FGLS) and Helpman Melitz and Rubinstein (2008) procedure.

a) Panel regressions[5]

Until 1990s, it has been a usual practice to estimate gravity equations using cross-section data. However, this type of estimation does not control for heterogeneity among countries. Consequently, results may vary substantially depending on the countries selected, leading to an estimation bias. To mitigate this problem, researchers have turned towards panel data, that is, cross-section gravity models for several consecutive years (Egger 2000, Rose and van Wincoop 2001, Mátyás 1998, Wall 2000, Egger and Pfaffermayr 2003, 2004; Glick and Rose 2002; Brun, Carrere, and de Melo 2002, Melitz 2007. See table A1 in the appendix).

Among the advantages of using a panel framework, we could cite that it allows to recognize how the relevant variables evolve through time and to identify the specific time or country effects (institutional, economical, cultural time-invariant or population-invariant factors). Additionally, the problem of potential multicollinearity that sometimes arises from cross-section data is completely avoided with panel data.

Fixed effect models assume that the unobserved heterogeneous component in the regression is constant over time. Dummies for importer and exporter are included in the sample, excluding one country to avoid perfect collinearity. As Mátyás (1997) points out, there may be a business cycle effect, which is common for all countries, but differs from one year to another. A dummy capturing this *time fixed effect* is then included.

However, some aspects affecting trade are not fixed along time, which may provoke a bias in the estimation. Consequently, Ruiz and Vilarrubia 2007 suggest including *exporter-yearly and importer-yearly dummies* in the regression. These variables absorb all country-specific factors, including those that vary over time.

It is probable that specific bilateral characteristics of partners influence trade –like remoteness. This can be controlled by including *country-pair fixed effect*. Again, those

---

[4]The criteria to compare between different methods are not always the same. The most common are the bias and the expected loss. Martinez-Zarzoso et al. (2008) construct a loss function that consists in the absolute error loss, defined as:

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

Other functions, such as the squared error loss, are also suggested.

[5] See the appendix for further information

effects can be time-varying or time-invariant; hence, two set of different dummies will be required.

Baltagi, Egger and Pfaffermayr (2003) classify the fixed effects into two groups: *main and interaction effects*. The first term makes reference to the usual fixed exporter, importer and time effects, whereas the second includes three types of dummies: one to control for country-pair fixed effects, another to control for exporter specific time-varying effects, and a last one in order to capture the same factors but from the importer's perspective.

The use of fixed effects also presents some **problems**. The most important disadvantage is related with dimension: the introduction of country specific or country pair dummies implies high computational costs. For that reason, Ruiz and Vilarrubia (2007) implement a regression with country triennial dummies, instead of country year dummies. They also make a test with country quinquennial dummies, but they obtain better results with triennial.

Additionally, any explanatory variables that do not vary across time in each country (or pair of countries in the case of country-pair fixed effects) will be perfectly collinear with the fixed effects, and should be dropped from the model. Then, country-pair fixed effect takes out of the gravity equation some important variables such as land area, common language, common borders or distance, and consequently, the effect of these variables on bilateral trade cannot be estimated.

Some authors have opted to assume that the unobserved component of the regression is distributed *randomly*. The difference between fixed and random effects is given by the correlation of the regressors. Fixed effects allow for correlation between the individual effects and the regressors ($Cov(\alpha_i, x_i) \neq 0$), whereas random effects impose that correlation to be zero. Orthogonality of the individual effects and the regressors is then required in order to use random effects. In other words, when assuming random effects we are implicitly assuming that the distribution of the unobserved heterogeneous component is distributed as a random variable with given mean and variance. In that case, there are specific factors in a country that affect trade but are not related with GDP, distance or other factors included as regressors. If we had enough evidence to suspect that the correlation is zero, we should employ random effects, because it provides more efficient estimators. However, if we are not sure of that uncorrelation, fixed effects should be preferred, since it is a less restrictive assumption: if we assume that the heterogeneity is better modeled with fixed effects and it is not, the equation is still consistent (though not efficient). In the reverse case, consistency cannot be assured.

Fratianni and Hoon Oh (2007) and Kavallari et al. (2008) are two examples of articles including both, fixed and random effects. They compare both models and applies different test in order to choose one of the two models: Breusch-Pagan test, LM test and Hausman test. Their results show that random effect model is preferred.

b) Poisson Pseudo Maximum Likelihood (PPML).

Another problem that arises in the estimation of the gravity equation is the "log or not to log" dilemma. It seems that the log-linearization of the error term changes the property of this error term and thus conduces to inefficient estimations due to heteroscedasticity. If data are homoscedastic, the variance of the error term is constant and its expected value is constant too. But if data are heteroscedastic, (as usual in trade data) the expected value of the error term is a function of the regressors. Then the conditional distribution of the dependent variable is altered and OLS is not efficient.

This point has been remarked several times by Silva and Tenreyro (Tenreyro, 2007; Silva and Tenreyro 2006, 2008). The essential point is that "*the log linearization of the empirical model in the presence of heteroscedasticity leads to inconsistent estimates because the expected value of the logarithm of a random variable depends on higher-order moments of its distribution*" (Silva and Tenreyro, 2006, p. 653). In the standard gravity equation

$$\ln(X_{ij}) = \ln\alpha_0 + \alpha_1 \ln(GDP_i) + \alpha_2 \ln(GDP_j) + \alpha_3 \ln(DIST_{ij}) + \ln(\varepsilon_{ij})$$

The expected value of the log-linearised equation would be:

$$E[\ln(X_{ij})] = E[\ln\alpha_0 + \alpha_1 \ln(GDP_i) + \alpha_2 \ln(GDP_j) + \alpha_3 \ln(DIST_{ij}) + \ln(\varepsilon_{ij})] =$$

$$E[\ln X_{ij}] = E[\ln(\alpha_0)] + \alpha_1 E[\ln(GDP_i)] + \alpha_2 E[\ln(GDP_j)] + \alpha_3 E[\ln(DIST_{ij})] + E[\ln(\varepsilon_{ij})]$$

Since $\ln E[\varepsilon_{ij}] \neq E[\ln(\varepsilon_{ij})]$ (Jensen's inequality), the conditional distribution of X_{ij} is altered and the estimation through OLS will result in misleading estimates. Heteroskedasticity does not affect the parameter estimates: the coefficients should be unbiased. However, it does bias the **variance** of the estimated parameters. As a result, the t-values for the estimated coefficients cannot be trusted.

The **source** of heteroscedasticity in data is not unique: the variance of the error term may vary with the regressors, with the dependent variable or with some other variable that has been omitted. In the gravity equation context, Kalirajan (2008) states that Anderson (1979) included in his theoretical model the *economic distance* between two countries. However, the common practice is to replace this concept by the *geographical distance*. The cost of this simplification is the omission of some important variables related with economic distance but not with geographical distance. Additionally, these non-included variables may be correlated with the included explanatory variables, so omitting them affects its variance, which will contain an upward bias. Among the aspects that are not easily quantifiable, Kalirajan cites the followings: *"large government size, weak and inefficient institution in home and partner countries in terms of, e.g. custom and regulatory environments, port efficiency and e-business and political influences through powerful lobbying by organized interest groups"* (Kalirajan, 2008, p. 1038). He claims that this imprecision in the measurement of distance leads to heteroscedastic error terms. This aspect has been also remarked by Silva and Tenreyro (2006), who point out that probably the variance of the error term is correlated with the countries' GDP and with the measure of distance.

The **solution** proposed by Silva and Tenreyro is to estimate the model in levels, instead of taking logarithms. In that case, OLS problems are avoided. They suggest two alternative methods: Nonlinear Least Squares (NLS) and Poisson Pseudo Maximum Likelihood (PPML), but finally show PPML as preferred. The reason is that NLS gives more weight to noisier observations, reducing henceforth the efficiency of the estimator.

4. Comparing empirical models and data

a) General model

This paper focuses on the impact of exchange-rate variables on trade flows, and on the comparison of different estimation methods. To this end, we use in a first step a baseline equation based on Anderson and van Wincoop (2003):

$$lnX_{ijt} = \alpha_1 lnGDP_{it} + \alpha_2 lnGDP_{jt} + \alpha_3 CONTIG_{ij} + \alpha_4 COMLA_{ij} + \alpha_5 COMCOL_{ij} + \alpha_6 COL45_{ij} + \alpha_7 SMCTRY_{ij} + \alpha_8 lnDIST_{ij} + \beta_i + \beta_j + \beta_t + \varepsilon_{ijt}$$

The dependent variable is the volume of exports in constant dollars (trade data from the CHELEM-CEPII database, price indexes from the World Bank and the IMF) from country j to i. $lnGDP_{it}$ and $lnGDP_{it}$ are the logarithms of real PPP-converted GDPs in each country. Next five variables are dummy variables. $DIST_{ij}$ is a variable representing the geodesic distance between i and j and is obtained from CEPII database. β_i is a vector of fixed effects for the exporting countries. β_j is a vector of fixed effects for the importing countries. β_t is a vector of fixed effects for time (yearly frequency).

Our sample includes 47 countries, of which all the countries of the EU15 and the CEE new European members, and 6 MENA countries (Morocco, Tunisia, Egypt, Turkey, Israel, Algeria). The time sample spans from 1980 to 2002. Hence, the total possible number of observations is 49,726. Due to missing data, the available number of observations is reduced to 34,457.

b) Results

The empirical model is estimated through different estimation methods: OLS, panel regression with fixed and random effect and simple and panel Poisson methodology.

If the true model is fixed effects, then OLS estimation yields biased and inconsistent estimates, since it deletes the individual dummies. Hence, we should, first of all, test the existence of fixed effects. We have tested it using a LR and LM on time and individual effects with the command xttest0 and xtreg, mle in Stata. Additionally, the output for the fixed effect regression includes the standard F-test for the joint significance of individual and time dummies. In all cases, we reject the null hypothesis of no fixed effects. Then, we assume that OLS provides biased and inconsistent estimates and we do not report those results.

In reference to the fixed effect estimation, we should mention that we are estimating a three-way error component model. Within transformation in Stata only deals with one of the three effects that we want to estimate. For that reason, we use the “fixed effect least square dummy variable” approach (Andrews et al. 2006); that is, we apply the within transformation to eliminate the exporter effects and introduce dummy variables to include the importer and time effects.

Another important aspect is whether we should assume fixed or random effect. This can be easily computed with a Hausman test. Under the null hypothesis, the random effect model is assumed to be consistent and efficient. In all cases, we reject the null, so random effect results (GLS estimation) are no efficient and we prefer the fixed effect estimation.

Finally, we suspect the data to be heteroscedastic. In order to check it, we implement the White's general test in OLS regressions and the Modified Wald statistic for groupwise heteroskedasticity in fixed effect models. In all our results, the null hypothesis of homoscedasticity is rejected; hence, the problem of heteroscedasticity should be taken into account. A simple method to correct it is to use robust standard errors: OLS and panel regressions assume that errors are both independent and identically distributed; robust standard errors relax either or both of those assumptions. Since Stata includes an option to for estimating robust standard errors, we simply use this option in each of the regression.

As expected, the exporter and importer real GDP both increase exports regardless to the estimation methods used. The distance also reduces exports though the elasticity is lower when using Poisson techniques. The estimated coefficients for GDP are near to 1, which is the expected order of magnitude, and the distance coefficient is also near to minus 1. Other gravity variables are also highly significant, and proximity (either in history or in space) tends to increase exports. The only exception is contiguity, which unexpectedly bears a negative sign when the gravity equation is estimated with panel with fixed effects while it displays the positive expected sign when Poisson is used. However, this variable is potentially collinear to the adjacency variable (close countries have a higher probability to share the same language), which could explain the sign of the estimate. The techniques of estimation seem to affect the magnitude of the parameters but not the sign for the other gravity variables. In particular, the impact of distance is found to be smaller under Poisson as in Martinez-Zarzoso et al. (2007). Unlike these authors, we do not appreciate important asymmetries in the coefficients of importer's and exporter's GDP using Poisson.

To have a first idea of the goodness of fit, we plot predicted over real value of exports for different techniques and compare the dispersions of the results. Graphics can be found in the appendix.

5. Impact of EMU on exports

a) Model

To assess the impact of exchange-rate variables and EMU on trade flows, we modify the previous equation, including some additional variables:

$$lnX_{ijt} = \alpha_1 lnGDP_{it} + \alpha_2 lnGDP_{jt} + \alpha_3 CONTIG_{ij} + \alpha_4 COMLA_{ij} + \alpha_5 COMCOL_{ij} + \alpha_6 COL45_{ij} + \alpha_7 SMCTRY_{ij} + \alpha_8 lnDIST_{ij} + \alpha_9 lnRER_{ijt} + \alpha_{10} VOL_{ijt} + \alpha_{11} RTAone + \alpha_{12} RTAboth + \alpha_{13} EMUone + \alpha_{14} EMUboth + \beta_i + \beta_j + \beta_t + \varepsilon_{ijt}$$

RER_{ijt} is the real exchange rate, computed using CPI and defined as the relative price of j to i (an increase therefore signals a real depreciation of the currency of country i compare to j).

VOL_{ijt} is a measure of volatility. This measure is one of the less obvious to build, as can be seen from the large number of volatility proxies that are available for the exchange rate. First of all, a large part of the financial literature highlights the fact that, as long as agents are information-seeking, only the unexpected part of exchange-rate volatility can have potential consequences on economic decisions. This is the reason why this literature has developed econometric models of the exchange-rate volatility (see e.g. ARCH models - and their various derivatives - for exchange rate series) aiming at extracting information from volatility series, and therefore allowing build unexpected volatility series.

In the longer run, exchange-rate are often described as following a random walk, and their standard deviation (or their coefficient of variation) is often enough to describe their volatility. While this might be true for nominal exchange rates, it is less relevant for real exchange rates, which are driven by fundamentals. In order to correctly measure their volatility, de-meaning is usually necessary, and a better measure of volatility is therefore the standard deviation of the rate of change of exchange-rate series.

We chose to use this last definition of volatility, applying it alternatively to monthly nominal and real exchange rates[6].

The definition of exchange rate volatility is therefore the following:

$$\text{VOL}=\sqrt{\text{Var}(\ln\text{ER}_{\text{ij}\tau} - \ln ER_{\text{ij},\tau\text{-1}})_{\{\tau=1\to12\}}}$$

Where $\text{ER}_{\text{ij}\tau}$ is the exchange rate, either nominal or CPI-deflated, and τ is monthly. Hence, we compute the volatility of the monthly exchange rate for a given year.

The impact of real exchange-rate changes on trade is now being quite well identified: a real appreciation usually has a deleterious impact on exports through a demand effect (lower competitiveness) or a supply effect (higher profitability of the traded goods sector compared to the non-traded goods sector).

The link between exchange-rate volatility and trade flows is less clear. According to McKenzie (1999), the elasticity of trade flows to exchange-rate volatility can be either positive or negative, and the results depend on the precise measure of volatility, on the estimation technique and on the sectors and countries concerned. Moreover, the impact of exchange-rate volatility might differ according to the countries under study: Sauer and Bohara (2001) show that exchange-rate volatility has a negative impact on African and Latin American exports, a non-significant impact on Asian exports and on developed countries exports.

Finally, a set of dummies is introduced. EMUone and EMUboth are two variables that take value one when one or both countries respectively belong to the EMU, which allow us to assess the effect of the EMU on non EU countries exports. Analogously, RTAone and RTAboth capture the effect of belonging to a regional trade agreement. With the inclusion of these variables, we intend to capture possible creation or diversion effects. Hence, a positive sign on these variables would imply that belonging to an RTA or to the EMU generate an increase in that country's exports, whereas a negative sign would mean that a diversion effect is taking place. Furthermore, we introduce two additional dummies (EMUimp and EMUexp) that distinguish among the cases in which only the exporter or only the importer belongs to the EMU.

[6] Notice that working on shorter-run data would call for the use of ARCH models. However, ARCH effects are usually shown to be less prevalent in the longer run (from the quarter to the year).

b) Results

Gains from anchoring to one money are assumed to be larger when the elasticity of trade to exchange rate volatility is higher, and this assumption allows investigating the potential gains of joining the euro area. Our database includes countries from different regions (MENA countries, Asian countries, New Member States in the European Union, other developed countries). We first study the impact of real exchange rate and volatility of the exchange rate on trade flows. By including dummies for trade and monetary agreements, we will try to estimate how a fixed peg could affect third countries exports to the EMU and to other zones.

When studying the impact of real exchange rate and volatility on exports (Table B2), the real exchange rate has the expected positive sign except in the Poisson regression with random effects. Elasticities differ from an estimation technique to another. In particular when individual fixed effect are used the elasticity is larger; a 10% depreciation leads to a 3% increase in bilateral exports while the effect is half lower when controlling for country pair fixed effects or using Poisson. This is a rather sensible price-elasticity estimate (working on the G7 countries, and relying on time-series econometrics, Hooper et al., 1998, find the long-run price-elasticity of exports to be ranging between .2 and 1.6).

The volatility of the exchange rate also has a detrimental effect on exports, which is significant at the 1% level. Here, a 10 point increase in volatility leads to a decrease between 7 and 8 % in exports according to panel estimations and more than 30% according to Poisson estimates.

Summing-up the whole-sample estimates, it appears that nominal or real exchange rate volatility is unambiguously detrimental to trade. As in Westerlund and Wilhelmson (2006) we found that the Poisson ML estimates are typically larger than their OLS counterparts.

Concerning the effect of RTA (Table B3), these agreements increase exports in all regressions when the exporter and the importer are members. The effect is also positive when only the importer or the exporter is a member of a RTA but this result is less robust in Poisson estimations.

Concerning the effect of the EMU, our results tend to show that the effects on exports are small or negative when significant. Poisson estimations support in general the most pessimistic views.

When we take into account the fact that the member of the Euro zone is the exporter or the importer (Table B4), our result do not support any diversion effect. On the opposite EMU seems to strengthen more imports from third countries than from EMU. Though in this model, export among EMU members appears negative. Poisson estimates indicate that both RTA and EMU have diversion effects since exporting to one member will reduce export of the third country exporter while fixed effects estimations drive opposite conclusions.

Our results are not so enthusiastic as previous results from Rose (2000) or from Micco et al. (2003) concerning the effect of EMU on trade among the members. Though these authors do not control for the exchange rate volatility, and Rose uses a cross section among a very large sample while Micco et al. (2003) use panel with country pair fixed effects for only 22 developed countries. These differences may explain the difference in the results. Micco et al. (2003) also found that EMU could have boosted trade with non-members.

6. Concluding remarks

The choice of an exchange rate regime, the possibility to peg or not to peg to the Euro and the effect the Euro could have on trade among members or with third countries or, lastly, between these third countries are puzzling questions.

First of all, our study confirms that the estimation technique is not neutral to study the effect of exchange-rate regimes – defined by the level of and the volatility of the (real) exchange rate – on exports, though it does not have so much importance for a basic model of trade flows. Different techniques lead to divergent results when the impact of EMU is studied. All in all, our results do not show strong diversion effects of the EMU.

This work could be extended in various directions. First, to conclude seriously about the appropriate estimations techniques some complementary tests should be performed and Monte Carlo simulations could also be used. Secondly, concerning the effect of the Euro on other countries exports, the period under study should be longer to capture the period after the Euro. Some asymmetries among members should be investigated and the effect of the Euro for trade among third countries could be an interesting issue to study.

7. References

Anderson, J. and Eric van Wincoop. 2003. Gravity with Gravitas: A Solution to the Border Puzzle, *American Economic Review*, 93: 170-192.

Anderson, J. E. 1979. "A theoretical foundation for the gravity equation". *American Economic Review*, 69: 106-116.

Andrews M., Schank T., Upward R. 2006. "Practical fixed effects estimation methods for the three-way error components model". *The Stata Journal*, 6 (4): 461–481

Baier, S., and Bergstrand, J. 2006. "Bonus vetus OLS: A simple approach for addressing the border-puzzle and other gravity issues', mimeo, March 2006

Baldwin, Richard and Virginia DiNino. 2006. "Euros and zeros: The common currency effect on trade in new goods", *HEI Working Paper* 21/2006

Baldwin et al. 2008. "Study on the impact of the Euro on Trade and FDI", *European Comission,* Economic Paper 321

Bergstrand, Jeffrey H. 1990. "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-industry Trade." Economic Journal 100: 1216-29.

Brun, Jean-François, Céline Carrère, Patrick Guillaumont and Jaime de Melo. 2002. "Has Distance Died? Evidence from a Panel Gravity Model," *CERDI* WP nº 2002/15

Burger, Martijn J., Frank G. van Oort and Gert-Jan M. Linders. 2009. "On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-Inflated Estimation", Forthcoming in *Spatial Economic Analysis*. Working paper version: ERIM Research Report 2009-003-ORG.

Cafiso, Gianluca. 2008. "The Euro's influence upon trade: Rose effect versus Border effect", *European Central Bank*, Working Paper 941.

Clark, Peter B., 1973, Uncertainty, Exchange Risk, and the Level of International Trade, *Western Economic Journal*, September, pp. 302-13.

Deardorff, Alan V. 1998. "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" In The Regionalization of the World Economy, edited by Jeffrey A. Frankel. Chicago: Univ. Ghicago Press (for NBER.

De Benedictis, Lucas and Claudio Vicarelli. 2005. "Trade Potentials in Gravity Panel Data Models," *Topics in Economic Analysis and Policy,* Berkeley Electronic Press, 5(1): 1386-1386.

Dixit, A.; Stiglitz, J. 1977. Monopolistic competition and optimum product diversity; *American Economic Review*, 67 (3), 297-308.

Egger, Peter. 2000. "A Note on the Proper Econometric Specification of the Gravity Equation," *WIFO Working Papers* 108.

Egger, Peter, and Michael Pfaffermayr. 2003. The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects. *Empirical Economics*, 28

Egger, Peter and Michael Pfaffermayr. 2004. "Distance, Trade and FDI: A SUR Hausman-Taylor Approach", *Journal of Applied Econometrics* 19 (2): 227-246

Evenett, Simon J. and Wolfgang Keller. 2002. "On Theories Explaining the Success of the Gravity Equation". *The Journal of Political Econom*y 110(2): 281-316.

Fratianni, Michele. 2007. "The gravity equation in international trade". *Handbook of international trade*, 2nd edition, Oxford University Press, Forthcoming

Fratianni, Michele and Chang Hoon Oh. 2007. "On the relationship between RTA expansion and Openness" Kelley School of Business.

Feenstra, Robert. 2002. "Border effects and the gravity equation. Consistent methods for estimation", *Scottish Journal of Political Economy*, 49(5): 491-506.

Freund, C. L. and Weinhold, D. 2004. The effect of the Internet on international trade*, Journal of International Economics 62, 171-189.*

Frankel, Jeffrey A., Ernesto Stein, Shang-Jin Wei. 1995. Trading blocs and the Americas: The natural, the unnatural, and the super-natural *Journal of Development Economics*, 47: 61-95.

Frankel, Jeffrey A., Ernesto Stein, Shang-Jin Wei, and Institute for International Economics. 1997. "Regional Trading Blocs in the World Economic System". *Washington DC: Institute for International Economics.*

Feenstra, Robert, James Markusen and Andrew Rose. 2001. "Using gravity equation to differentiate among alternative theories of trade" *Canadian Journal of Economics*, 34 (2)

Gil-Pareja, Salvador, Rafael Llorca-Vivero, José Martínez-Serrano, Josep Oliver-Alonsol. 2005. "The Border Effect in Spain", *The World Economy*, 28 (11):1617-1631

Gil-Pareja, Salvador, Rafael Llorca-Vivero and José Martínez-Serrano. 2008. " Trade effects of monetary agreements: Evidence for OECD countries", *European Economic Review*, 52 (4):733-755

Gil-Pareja, Salvador, Rafael Llorca-Vivero, José Martínez-Serrano.2009. "Continental bias in trade", *Fundación de las Cajas de Ahorros,* Documento de trabajo nº 440/2009

Glick, Reuven and Andrew Rose. 2002. "Does a currency union affect trade? The time-series evidence," *European Economic Review*, 46(6): 1125-1151.

Head, Keith and Thierry Mayer, 2000. "Non-Europe: The magnitude and causes of market fragmentation in the EU," *Review of World Economics*, 127(2): 284-314.

Helpman, Elhanan and Krugman, Paul. 1985. "Market Structure and Foreign Trade". Cambridge, MA: MIT Press.

Henderson, Daniel and Daniel Millimet. 2008. “Is gravity linear?”, *Journal of Applied Economics* 23: 137–172

Hooper, Peter, and Steven Kohlhagen, 1978, The Effect of Exchange Rate Uncertainty on the Prices and Volume of International Trade, *Journal of International Economics*, 8: 483-511.

Kavallari, A.; Maas, S. and Schmitz, P.M., 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 44217

Kang, Heejoon and Michele Fratianni. 2006. "International Trade Efficiency, the Gravity Equation, and the Stochastic Frontier", *Kelley School of Business*, WP 2006-08

Kimura, Fukunari and Hyun-Hoon Lee. 2006. “The gravity equation in International Trade in Services”, Review of World Economics, 142 (1): 92-121

Lahrèche-Révil, Amina and Juliette Milgram, 2006. "Exchange-rate policies and trade in the MENA countries," ThE Papers 06/07,*University of Granada*.

Leamer, Edward and Robert. 1971. *Journal of International Economics*, 1(3): 359-361

Stern Quantitative international economics:

Linders, Gert-Jand and Henri L.F. de Groot. 2006. "Estimation of the gravity equation in the presence of zero flows", *Tinbergen Institute Discussion Paper*, 072/3

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

Martin and Pham. 2008, "Estimating the Gravity Equation when Zero Trade Flows are Frequent", MPRA Paper No. 9453

Márquez-Ramos, Laura; Inmaculada Martínez-Zarzoso and Celestino Suárez-Burguet. 2007. "The Role of Distance in Gravity Regressions: Is There Really a Missing Globalisation Puzzle?", *The B.E. Journal of Economic Analysis & Policy*, 7(1)

Martinez-Zarzoso Inmaculada and F. Nowak-Lehman. 2003. Augmented Gravity Model: An Empirical Application to Mercosur-European Union Trade Flows, *Journal of Applied Economics*, 6 (2): 291-316.

Martínez Zarzoso, Inmaculada and Laura Márquez Ramos. 2005. "International Trade, Technological Innovation And Income: A Gravity Model Approach," Instituto Valenciano de Investigaciones Económicas, S.A, Working Papers. Serie EC 2005-15.

Martínez-Zarzoso, Inmaculada; Felicitas Nowak-Lehmann D. and Sebastian Vollmer et al. 2007. "The log of gravity revisited". *Center for European, Governance and Economic Development Research.* Discussion Papers 64.

Melitz, Marc. 2003. "The impact of trade on intraindustry reallocations and aggregate industry productivity" *Econometrica* 71, 1695--1725

Melitz, Marc. 2007. "North, South and Distance in the gravity equation", *European Economic Review* 51: 971--991

Melitz, Marc and Giancarlo Ottaviano. 2008. "Market Size, Trade and Productivity", *Review of Economic Studies* (75): 295—316

Alejandro Micco, Alejandro, Ernesto Stein and Guillermo Ordoñez. 2003. "The currency union effect on trade: early evidence from EMU," *Economic Policy*, 18(37), pages 315-356.

Rose, Andrew K. 2000. "One Money, One Market: The Effect of Common Currencies on Trade" *Economic Policy*, 15(30): 9.

Rose, Andrew and Eric van Wincoop. 2001. "National Money as a Barrier to International Trade: The Real Case for Currency Union" *American Economic Review,* 91(2): 386-390

Sapir, André. 2001. "Domino effects in Western European Regional Trade" *European Journal of Political Economy,* 17: 377-388

Serlenga, Laura and Yongcheol Shin. 2007. "Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors" *Journal of Applied Economics 22:* 361-381

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

Silva, J. M. C. S. and Silvana Tenreyro. 2008. "Trading Partners and Trading Volume: Implementing the Helpman-Melitz-Rubinstein Model Empirically" University of Essex, Discussion Paper Series 662.

Soloaga, I. and Winters, A. 2001. "Regionalism in the nineties: what effect on trade?" *North American Journal of Economics and Finance*, 12, 1-29.

Tenreyro, Silvana. 2007. "On the Trade Impact of Nominal Exchange Rate Volatility" *Journal of Development Economics*, 82(2): 485-508.

Wei, Shang-Jin. 1996. "Intra-National Versus International Trade: How Stubborn Are Nations in Global Integration?" *National Bureau of Economic Research* WP 5531.

Westerlund, Joakim and Fredrik Wilhelmsson. 2006. "Estimating the gravity model without gravity using panel data" Nationalekonomiska institutionen Department of Economics, WP

Wolf, H. C. 1997. "Patterns of Intra- and Inter-State Trade". *National Bureau of Economic Research* Working Paper 5939.

Wong, Wei-Kang. 2008. "Comparing the fit of the gravity model for different cross-border flows", *Economic letters* 99: 474-477

8. Appendix

A. Literature review

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

Article	Effects included	Disaggregation level	M, X or T
Matyas (1997)	Fixed importer, exporter and time effects	11 countries, 1982-1994	Exports
Rose and van Wincoop (2001)	Time effects, country specific fixed effects	Data at five-year intervals between 1970 and 1995 covering almost 200 countries	Bilateral trade
Glick and Rose (2002)	- Country pair fixed effects - They impose the restriction that the country-pair effects are symmetric (i.e., αij = αji).	Panel data set covering 217 countries from 1948 through 1997	Real bilateral trade
Baltagi, Egger and Pfaffermayr (2003)	- Fixed importer, exporter and time effects - Country pair fixed effects - Importer-time effect - Exporter time effect for exporter specific time variant effects	Panel of bilateral trade between the triad (EU15, USA and Japan) economies and their 57 most important trading partners over the period 1986–1997	Real bilateral exports
Micco, Stein and Ordoñez (2003)	- Time effects - Country pair fixed effects - No individual effects	22 developed countries; 1992 - 2002	Bilateral trade (sum of imports and exports
De Benedictis and Vicarelli (2004)	- Bilateral (country-pair) fixed effects - Dynamic effects (Arellano and Bond estimator)	Export equation for each of former 11 Eurozone countries to 32 importer countries; period 1991-2000.	Exports
Cheng and Wall (2005)	- Country-pair fixed effects - Time effects	Balanced panel with 3,188 observations (797 unidirectional country pairs in each of four years: 1982, 1987, 1992, and 1997	Real exports
Fratianni and Hoon Oh (2007)	- Country pair and time fixed effects - Random effects	143 countries for the period 1980-2003.	Real bilateral imports
Cafiso (2008)	Country pair and time fixed effects	Manufacture export between 24 OECD countries (sectors 15-37, ISIC Rev. 3); 1993-2003	Exports
Fidrmuc (2008)	Country pair, time effects	19 OECD countries between 1980 and 2002.	Bilateral trade flows (average of exports and imports)
Henderson and Millimet (2008)	Country specific, country pair fixed effects	1993 and 1997; US data. 25 two-digit SIC industries;	Nominal value of exports
Hoon Oh and Travis Selmier II (2008)	- Time-invariant country-pair fixed effects - Time-invariant country-pair random effects	1980–2001, 10,520 observations for 859 bilateral pairs	Imports
Kavallari et al. (2008)	Random effects	German imports of olive oil from 14 exporting countries; 1995-2006.	Imports
Ruiz and Vilarrubia (2008)	- Fixed importer, exporter and time effects - Exporter-period and importer-period dummies (annual, triennial and quinquennial)	205 countries from 1948 to 2005 (regression over the top 100 exporters)	Bilateral trade

Table A2. Alternative estimation methods in the literature to deal with the problem of zero-flows and heteroscedasticity.

Article	Countries and years	Estimation methods (preferred)	Disaggregation level	M, X or T	Simulation studies
Silva and Tenreyro (2006)	Cross section of 136 countries in 1990 (18,360 observations)	- **PPML** , NLS , GPML, OLS, ET-tobit , OLS(*y* >0,.5) OLS (y+1)	Aggregated data Dummies for FTAs	Bilateral trade flows	- PPML, NLS, GPML OLS; OLS(*y* + 1); truncated OLS ET-tobit. - Four different patterns of heteroscedasticity
Martínez-Zarzoso (2007)	3 datasets: 1) 180 countries; 1980-2000 2) 47 countries; 1980-1999 3) 65 countries; data for every 5 years over 1980-1999.	- **FGLS**,Gamma, Poisson, Heckman	Aggregated data	Exports	- OLS, NLS, Gamma Pseudo Maximum Likelihood (GPML), PPML and FGLS
Siliverstov and Schumacher (2007)	1988 to 1990; 22 OECD countries	OLS, PQML	Disaggregated data: 25 three-digit ISIC Rev.2 industries and the manufacturing as a whole	Average annual trade flows	No
Westerlund and Wilhelmsson (2007)	1992-2002; EU and other developed countries (35256 observations)	OLS, fixed effect PML	Aggregated data	Nominal imports	- OLS, truncated OLS, OLS (y+1), PPML - Two patterns of heteroscedasticity
Helpman, Melitz and Rubinstein (2008)	1970-1997; 158 countries	**HMR** (Probit and OLS), NLS, semiparametric, non-parametric	Aggregated data	Exports	No
Martin and Pham (2008)	Dataset from Silva and Tenreyro (2006): cross section of 136 countries in 1990 (18,360 observations)	- Truncated OLS, ET-Tobit, PPML, Heckman ML, Heckman 2SLS	Aggregated data	Bilateral trade	- Truncated OLS, OLS (y+1), truncated NLS, censored NLS, GPML, PPML, truncated PPML, ET Tobit, Poisson-Tobit,Heckman
Silva and Tenreyro (2008)	1986 (cross-section); 158 countries;	HMR (Probit and OLS), NLS, semiparametric, non-parametric, GPML	Aggregated data	Exports	No
Burger et al. (2009)	138 countries 1996-2000	OLS, Poisson and modified Poisson (negative binomial, zero-inflate: ZIPPML, NBPPML)	Aggregated data	Average of yearly exports	No

B. Estimation results

Table B1. Baseline model. Logarithm of exports

	Panel FELSDV	**Panel FE** (country pairs)	**Panel GLS**	**Poisson FELSDV**	**Poisson FE** (country pairs)	**Poisson GLS**
Log of exporter real GDP	0.769*** [0.036]	0.826*** [0.023]	0.536*** [0.046]	0.699*** [0.001]	0.640*** [0.001]	0.714*** [0.001]
Log of importer real GDP	0.577*** [0.039]	0.831*** [0.023]	0.237*** [0.007]	0.546*** [0.001]	0.657*** [0.001]	0.642*** [0.000]
Contiguity	-0.333*** [0.042]		0.489*** [0.047]	0.454*** [0.001]		0.619*** [0.000]
Common Language	0.491*** [0.025]		1.048*** [0.032]	0.226*** [0.000]		0.336*** [0.000]
Colony	0.349*** [0.035]		0.573*** [0.046]	-0.137*** [0.001]		0.073*** [0.001]
Common Colony	0.642*** [0.077]		-0.031 [0.072]	-0.273*** [0.001]		0.075*** [0.001]
Colony after 1945	0.666*** [0.051]		0.863*** [0.078]	0.244*** [0.001]		0.339*** [0.001]
Same Country	0.692*** [0.065]	-0.390*** [0.119]	0.091 [0.056]	0.170*** [0.001]	-0.332*** [0.003]	-0.036*** [0.001]
Log of Distance	-1.248*** [0.008]		-0.808*** [0.010]	-0.713*** [0.000]		-0.618*** [0.000]
Constant	-20.697*** [1.371]	-37.515*** [1.180]	-8.449*** [1.199]			-16.654*** [0.439]
Observations	38643	38643	38643	41950	41950	41950
R-squared	0.736	0.180	0.18			
-2 Log Likelihood				19706805,2	9485543,4	61277847.20
Exporter fixed effects	Yes (Within)	No	No	Yes (Within)	No	No
Importer fixed effects	Yes	No	No	Yes	No	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country pair fixed effects	No	Yes	No	No	Yes	No
Modified Wald test	0.00	0.00				
Hausman test			1.00			0.00

Notes: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable is logarithm of exports for panel regressions and exports in levels for poisson regressions.

Table B2. Effects of real exchange rate and volatility on exports.

	FELSDV	Panel FE (country pairs)	Panel GLS	Poisson FELSDV	Poisson FE (country pairs)	Poisson GLS
Log of exporter real GDP	0.897*** [0.046]	1.040*** [0.075]	0.978*** [0.048]	0.936*** [0.001]	0.824*** [0.001]	0.983*** [0.001]
Log of importer real GDP	0.797*** [0.048]	0.951*** [0.076]	0.971*** [0.006]	0.681*** [0.001]	0.835*** [0.001]	0.818*** [0.000]
Contiguity	-0.388*** [0.043]		-0.151*** [0.042]	0.384*** [0.001]		0.437*** [0.000]
Common Language	0.540*** [0.026]		0.958*** [0.027]	0.259*** [0.001]		0.351*** [0.000]
Colony	0.349*** [0.035]		0.282*** [0.033]	-0.103*** [0.001]		0.019*** [0.001]
Common Colony	0.368*** [0.076]		0.338*** [0.084]	-0.316*** [0.002]		0.171*** [0.001]
Colony after 1945	0.748*** [0.050]		0.616*** [0.057]	0.345*** [0.001]		0.187*** [0.001]
Same Country	0.914*** [0.067]	-0.380* [0.200]	0.633*** [0.055]	0.301*** [0.001]	-0.329*** [0.003]	0.321*** [0.001]
Log of Distance	-1.230*** [0.009]		-1.061*** [0.009]	-0.724*** [0.000]		-0.668*** [0.000]
Log of RERij	0.337*** [0.031]	0.148*** [0.048]	0.282*** [0.030]	0.136*** [0.001]	-0.024*** [0.001]	-0.199*** [0.001]
NER Volatility	-0.743*** [0.159]	-0.850*** [0.124]	-2.509*** [0.164]	-3.395*** [0.009]	-1.170*** [0.008]	-4.847*** [0.008]
Constant	-30.017*** [1.677]	-46.702*** [3.875]	-36.786*** [1.260]			-35.016*** [0.107]
Observations	34893	34893	34893	35324	35324	35324
R-squared	0.738	0.194	0.62			
-2 Log Likelihood				17477124.47	8257511.15	30120048
Exporter fixed effects	Yes (Within)	No	No	Yes (Within)	No	No
Importer fixed effects	Yes	No	No	Yes	No	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country pair fixed effects	No	Yes	No	No	Yes	No
Modified Wald test	0.00	0.00				
Hausman test			0.00			1.00

Notes: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable is logarithm of exports for panel regressions and exports in levels for poisson regressions.

Table B3. Effect of RTAs on exports

	FELSDV	Panel FE	Panel GLS	Poisson FELSDV	Poisson FE	Poisson GLS
Log of exporter real GDP	0.881*** [0.047]	1.071*** [0.077]	0.968*** [0.048]	0.969*** [0.001]	0.852*** [0.001]	1.012*** [0.001]
Log of importer real GDP	0.774*** [0.048]	0.979*** [0.078]	0.961*** [0.006]	0.701*** [0.001]	0.865*** [0.001]	0.824*** [0.000]
Contiguity	-0.372*** [0.042]		-0.094** [0.040]	0.352*** [0.001]		0.392*** [0.000]
Common Language	0.538*** [0.026]		0.997*** [0.027]	0.260*** [0.001]		0.378*** [0.000]
Colony	0.354*** [0.035]		0.274*** [0.033]	0.011*** [0.001]		0.084*** [0.001]
Common Colony	0.387*** [0.076]		0.578*** [0.082]	-0.336*** [0.002]		0.212*** [0.001]
Colony after 1945	0.734*** [0.049]		0.579*** [0.055]	0.299*** [0.001]		0.255*** [0.001]
Same Country	0.911*** [0.066]	-0.378* [0.198]	0.770*** [0.055]	0.363*** [0.001]	-0.329*** [0.003]	0.454*** [0.001]
Log of Distance	-1.225*** [0.009]		-0.954*** [0.010]	-0.632*** [0.000]		-0.566*** [0.000]
Log of RERij	0.339*** [0.030]	0.145*** [0.048]	0.261*** [0.029]	0.128*** [0.001]	-0.023*** [0.001]	-0.245*** [0.001]
NER Volatility	-0.692*** [0.160]	-0.729*** [0.120]	-2.138*** [0.168]	-2.988*** [0.009]	-1.203*** [0.008]	-4.075*** [0.008]
One partner has FTA	0.260*** [0.023]	0.092* [0.048]	0.666*** [0.023]	-0.001 [0.001]	-0.062*** [0.001]	0.216*** [0.000]
Both partners have FTA	0.143*** [0.020]	0.294*** [0.045]	0.437*** [0.020]	0.447*** [0.000]	0.146*** [0.000]	0.496*** [0.000]
One partner in EMU	0.187*** [0.024]	0.002 [0.036]	0.474*** [0.028]	-0.036*** [0.001]	0.008*** [0.001]	0.002*** [0.001]
Both partners in EMU	-0.345*** [0.045]	-0.049 [0.047]	-0.011 [0.041]	0.033*** [0.001]	-0.075*** [0.001]	0.094*** [0.001]
Constant	-29.159*** [1.705]	-48.279*** [3.978]	-37.553*** [1.267]			-36.963*** [0.109]
Observations	34893	34893	34893	35324	35324	35324
R-squared	0.740	0.199	0.64			
-2 Log Likelihood				16632031,12	8159163,15	28399956,0
Importer fixed effects	Yes	No	No	Yes	No	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country pair fixed effects	No	Yes	No	No	Yes	No
Modified Wald test	0.00	0.00				
Hausman test			0.00			0.00

Notes: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable is logarithm of exports for panel regressions and exports in levels for poisson regressions.

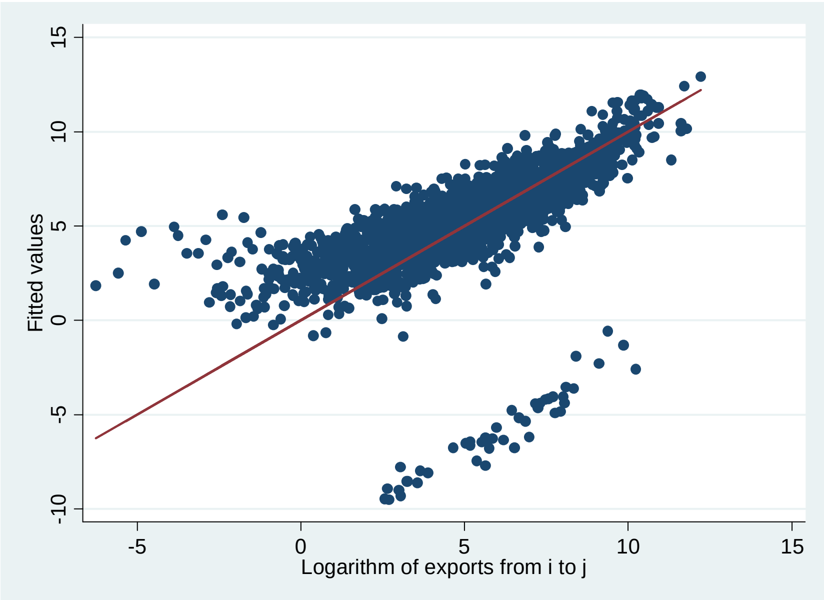
Table B4. RTA and EMU diversion effects

	FELSDV	Panel FE	Panel GLS	Poisson FELSDV	Poisson FE	Poisson GLS
Log of exporter real GDP	0.882***	1.070***	0.943***	0.971***	0.852***	1.018***
	[0.047]	[0.077]	[0.049]	[0.001]	[0.001]	[0.001]
Log of importer real GDP	0.773***	0.980***	0.959***	0.699***	0.865***	0.824***
	[0.049]	[0.078]	[0.006]	[0.001]	[0.001]	[0.000]
Contiguity	-0.372***		-0.094**	0.352***		0.393***
	[0.042]		[0.040]	[0.001]		[0.000]
Common Language	0.538***		0.999***	0.260***		0.377***
	[0.026]		[0.027]	[0.001]		[0.000]
Colony	0.354***		0.275***	0.011***		0.085***
	[0.035]		[0.034]	[0.001]		[0.001]
Common Colony	0.387***		0.583***	-0.336***		0.211***
	[0.076]		[0.082]	[0.002]		[0.001]
Colony after 1945	0.734***		0.580***	0.299***		0.254***
	[0.049]		[0.055]	[0.001]		[0.001]
Same Country	0.911***	-0.377*	0.780***	0.363***	-0.330***	0.454***
	[0.066]	[0.199]	[0.056]	[0.001]	[0.003]	[0.001]
Log of Distance	-1.225***		-0.952***	-0.632***		-0.567***
	[0.009]		[0.010]	[0.000]		[0.000]
Log of RERij	0.339***	0.148***	0.265***	0.125***	-0.019***	-0.248***
	[0.031]	[0.048]	[0.029]	[0.001]	[0.001]	[0.001]
NER Volatility	-0.692***	-0.728***	-2.108***	-2.986***	-1.204***	-4.101***
	[0.160]	[0.120]	[0.168]	[0.009]	[0.008]	[0.008]
One partner has FTA	0.260***	0.094*	0.652***	-0.001	-0.062***	0.219***
	[0.023]	[0.048]	[0.023]	[0.001]	[0.001]	[0.000]
Both partners have FTA	0.143***	0.297***	0.424***	0.447***	0.147***	0.497***
	[0.020]	[0.045]	[0.020]	[0.000]	[0.001]	[0.000]
EMUimp	0.181***	-0.069	0.661***	-0.047***	-0.053***	-0.026***
	[0.031]	[0.049]	[0.032]	[0.001]	[0.001]	[0.001]
EMUexp	0.193***	0.073*	0.247***	-0.027***	0.066***	0.034***
	[0.028]	[0.041]	[0.037]	[0.001]	[0.001]	[0.001]
Both partners in EMU	-0.158***	-0.047	0.426***	-0.003***	-0.067***	0.102***
	[0.048]	[0.047]	[0.044]	[0.001]	[0.001]	[0.001]
Constant	-29.161***	-48.289***	-36.873***			-37.100***
	[1.705]	[3.978]	[1.275]			[0.109]
Observations	34893	34893	34893	35324	35324	35324
R-squared	0.740	0.200	0.64			
-2 Log Likelihood				16631692,8	8141761,89	28396178,0
Exporter fixed effects	Yes (Within)	No	No	Yes (Within)	No	No
Importer fixed effects	Yes	No	No	Yes	No	No
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country pair fixed effects	No	Yes	No	No	Yes	No
Modified Wald test	0.00	0.00				
Hausman test			0.00			0.00

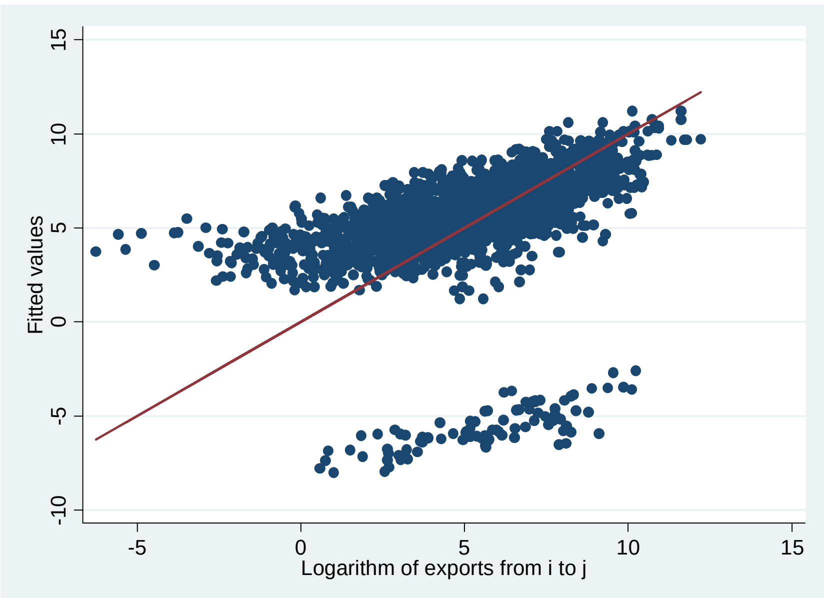
Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in brackets. EMUimp takes value one when the importer belongs to EMU and EMUexp when the exporter belongs to EMU. The dependent variable is logarithm of exports for panel regressions and exports in levels for poisson regressions.

C. Cross-validation for the different estimation methods in year 2002

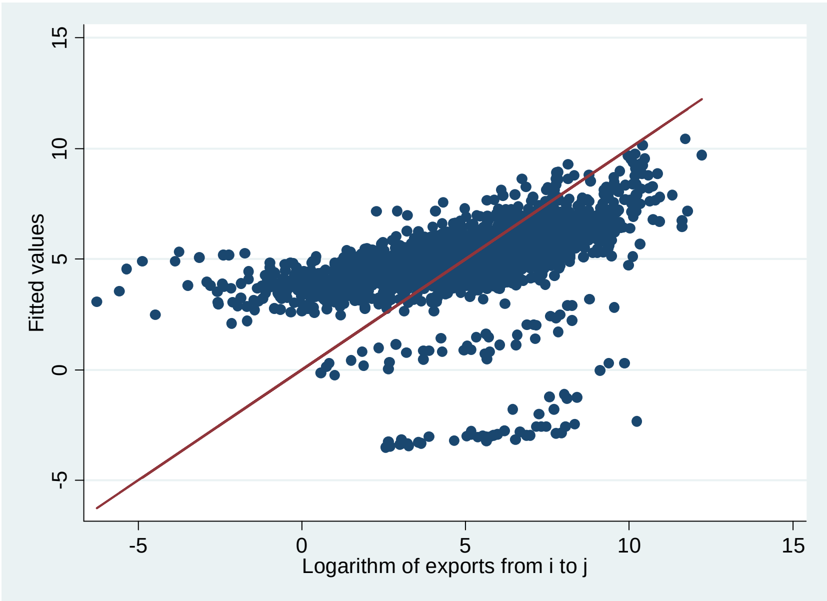
C1. Cross- validation for panel regression. Exporter, importer and time fixed effects



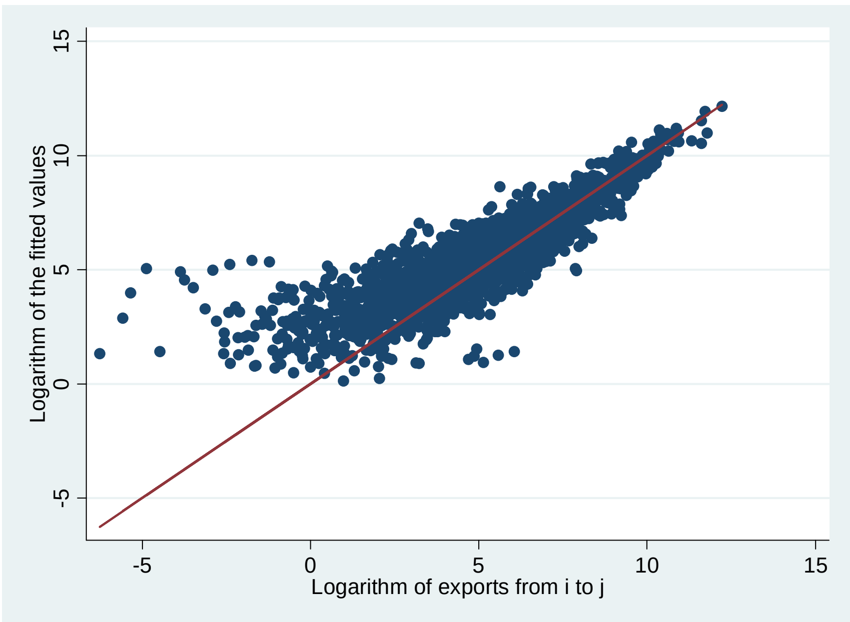
C2. Cross- validation for Panel regression. Country pair fixed and time effects



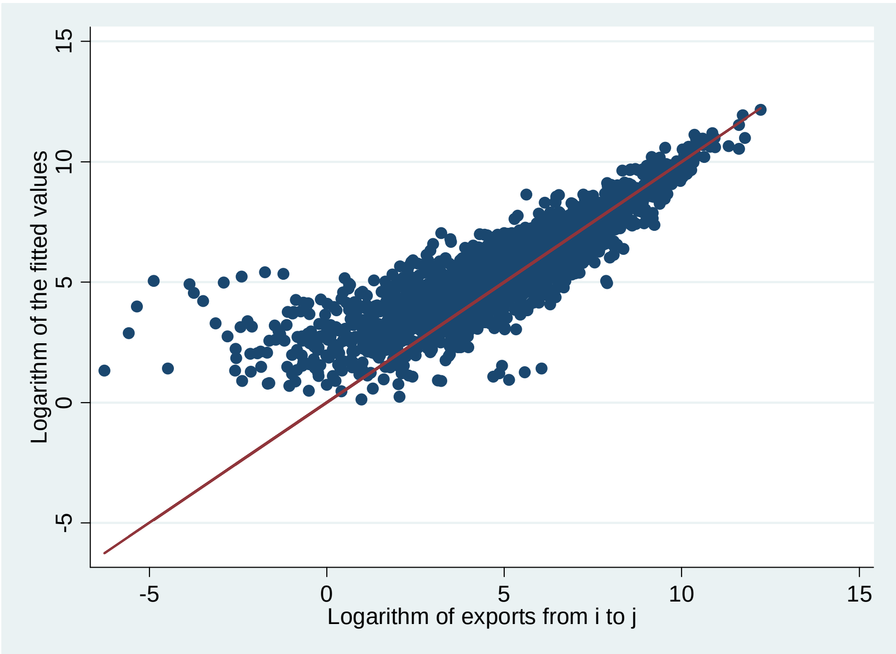
C3. Cross- validation for Panel regression. Random effects



C4. Cross- validation for Poisson regression. Exporter, importer and time fixed effects



C5. Cross-validation for Poisson regression. Country pair and time fixed effects



C6. Cross-validation for Poisson regression with random effects