The gravity model for international trade: Specification and estimation issues in the prevalence of zero flows

Krisztin, Tamás; Fischer, Manfred M.

The gravity model for international trade: Specification and estimation issues

Tamás Krisztin and Manfred M. Fischer

Institute of Economic Geography and GIScience

Vienna University of Economics and Business

Abstract The Poisson gravity model along with pseudo ML methods has become a popular way to model international trade flows. This approach has several econometric advantages that we outline in the paper. We argue that estimating the parameters by ML would only be justified statistically if the trade flows were independent. Such an assumption, however, is generally not valid, and a failure to account for spatial dependence may lead to biased parameter estimates and misleading inferences. To overcome this estimation problem we suggest eigenvector spatial filtering variants of the Poisson gravity model (without and with zero-inflation) along with pseudo ML estimation.

Keywords: Poisson gravity model, zero-inflation, origin-destination dependence, spatial filtering

JEL Classification: C13, C21, F10, R15

March 3, 2015

1 Introduction

The pioneering work of Tinbergen (1962) triggered a vast empirical literature on the gravity model for international trade. In its simplest form, the model states that the volume of trade between any two countries is proportional to the product of their gross domestic products and a distance deterrence function where distance is broadly construed to include all factors that might create trade resistance. The popularity of Tinbergen's log-linear specification of the gravity model is partly due its apparently good performance in modelling trade flows, and partly due to the strong theoretical foundations provided in papers such as Anderson (1979), and Anderson and van Wincoop (2003).

Despite the extensive use of the log-linear specification of the gravity model in empirical research, the log-linear specification of the gravity model along with OLS estimation is inappropriate for several reasons. First of all, bilateral trade is frequently zero and then log-linearisation is infeasible. Most of the studies estimate the log-linear model on samples of countries. By disregarding countries that do not trade with each other, these studies give up important information inherent in the data, and generate biased estimates as a result (Helpman et al. 2008). Second, Santos Silva and Tenreyro (2006) have shown that log-linearisation of the gravity model leads to inconsistent estimates in the presence of heteroscedasticity in trade levels. This is so, because the expected value of the logarithm of a random variable depends on log-linear moments of its distribution, and hence if the residuals of the multiplicative gravity model are heteroscedastic, the log-transformed residuals will be correlated with the covariates in general.

To overcome these econometric problems, Santos Silva and Tenreyro (2006) propose a Poisson specification of the gravity model along with the Poisson pseudo maximum likelihood (PPML) estimator introduced by Gourieroux et al. (1984). Santos Silva and Tenreyro (2006, 2010, 2011) provide simulation evidence that the PPML estimator is well behaved even when the conditional variance is far from being proportional to the conditional mean. Moreover, the fact that the dependent variable has a large proportion of zeros does not affect the performance of the estimator.

As a result, a number of empirical studies of trade flows have applied the PPML estimator (see Linders et al. 2008, Martin and Pham 2008, Burger et al. 2009, Liu 2009, Westerlund and Wilhelmsson 2011, Bosquet and Boulhol 2010, Martínez-Zarzoso 2013 among others). The negative binomial gravity model specification with the negative binomial pseudo maximum likelihood (NBPML) estimator has also received increasing attention in the trade literature (see Burger et al. 2009, Westerlund and Wilhelmsson 2011). The negative binomial distribution assumes that the conditional variance is a linear combination of the conditional mean and of its square. This estimator is attractive, but is scale dependent (as noticed by Bosquet and Boulhol 2010) and hence not appropriate when applied to continuous dependent variables, such as trade flows, for which the choice of the unit measure is arbitrary.

Estimating the parameters by PPML, however, would only be justified statistically if we believed that the trade flows were independent. Such an assumption is generally not valid, since flows are fundamentally spatial in nature. Spatial dependence is more likely than spatial independence. A failure to account for spatial dependence in trade flows may lead to biased parameter estimates and incorrect conclusions (Griffith and Fischer 2013). This estimation problem has been largely neglected so far. We propose a spatial filtering variant of the Poisson gravity model along with PPML to overcome this problem. A virtue of this model specification is that filtering out spatial dependence in trade flows reduces potential bias in the parameter estimates.

The remainder of the paper is organized as follows. Section 2 briefly describes the log-linear specification of the gravity model, followed by a discussion of the econometric problems raised by the estimation of this model specification. Section 3 shifts attention to pseudo maximum likelihood estimation techniques applied to the Poisson gravity model specification. In section 4 we present the zero-inflated extension of the Poisson (ZIP) gravity model along with PPML estimation. Section 5 continues to describe the approach of spatial filtering as applied to flow data, leading to spatial filtering variants of the Poisson gravity model (without and with zero-inflation).

In section 6 we use these model specifications to quantitatively assess the determinants of trade flows, uncovering significant differences in the role of distance measures from those predicted by the standard Poisson gravity model specification. The comparisons are performed using a real world example that covers a cross-section of 146 countries. The dataset consists of 21,170 observations on bilateral trade flows where 53% of the observations are zero. We use unidirectional export trade values, measured in terms of millions of US dollars, as an indicator of trade volume in 2000. Section 7 contains concluding remarks.

2 The traditional gravity equation

Let Y_{ij} denote the bilateral trade between countries $i = 1, ..., n$ and $j = 1, ..., n$ (with $i \neq j$), as measured by exports from country i to country j. For convenience, the total number of observations, which is given by $n(n-1)$, is denoted by N. In its simplest form, the gravity equation for trade from country i to country j, Y_{ij}, is proportional to the product of the two countries' gross domestic product (*gdp*), denoted by X_i and X_j, and a distance deterrence function (usually a power function) involving distance, D_{ij}, between i and j, broadly construed to include all factors that might create trade resistance:

$$Y_{ij} = \beta_0 X_i^{\beta_1} X_j^{\beta_2} D_{ij}^{\beta_3}. \tag{1}$$

β_0, β_1, β_2 and β_3 are unknown parameters. Typically, the stochastic version of this gravity equation has the form

$$Y_{ij} = \beta_0 X_i^{\beta_1} X_j^{\beta_2} D_{ij}^{\beta_3} \zeta_{ij} \tag{2}$$

where ζ_{ij} is a disturbance term with $\mathrm{E}\left[\zeta_{ij}|X_i, X_j, D_{ij}\right] = 1$, assumed to be statistically independent of the explanatory variables X_i, X_j and D_{ij}. This leads to

$$\mathrm{E}\left[Y_{ij}|X_i, X_j, D_{ij}\right] = \beta_0 X_i^{\beta_1} X_j^{\beta_2} D_{ij}^{\beta_3}. \tag{3}$$

The most prevalent approach to estimate the multiplicative gravity model for trade given by Eq. (2) is to use a log-log transformation yielding

$$\ln Y_{ij} = \ln \beta_0 + \beta_1 \ln X_i + \beta_2 \ln X_j + \beta_3 \ln D_{ij} + \ln \zeta_{ij} \tag{4}$$

and then to estimate the parameters of interest by ordinary least squares (OLS).

But this practice is inappropriate for a number of reasons. First of all, Y_{ij} can be zero and then log-linearisation is infeasible. Indeed, the level of trade between any two countries is frequently zero. About half of the observations on total bilateral trade in the datasets used by Santos Silva and Tenreyro (2006), Helpman et al. (2008), and Burger et al. (2009), for example, are zero trade flows. Second, even if all trade observations are strictly positive, it should be noted that the validity of the estimation approach critically depends on the assumption that ζ_{ij}, and hence $\ln \zeta_{ij}$, are statistically independent of the explanatory variables. Santos Silva and Tenreyro (2006) argue that, if we assume ζ_{ij} to follow a log-normal distribution, with $\mathrm{E}\left[\zeta_{ij}|X_i, X_j, D_{ij}\right] = 1$ and variance-covariance $\sigma^2_{ij} = f(X_i, X_j, D_{ij})$, then the log-linearised version of these disturbances has $\mathrm{E}\left[\ln \zeta_{ij}|X_i, X_j, D_{ij}\right] = -\frac{1}{2}\ln(1+\sigma^2_{ij})$, which exhibits dependence on the explanatory variables, and thus violates the condition for consistency of OLS. Note that heteroscedasticity is important not only for the efficiency of an estimator, but also for its consistency, because Eq. (4) produces the estimate of $\ln Y_{ij}$ rather than Y_{ij} itself, which leads to biased estimates due to Jensen's inequality.

Several methods have been suggested to deal with the zero flows problem (see, for example, Frankel 1997, pp.145-146). The approach followed by the large majority of studies excludes zero trade flows from the data set and estimates the parameters of interest on the truncated sample of countries that have only positive trade flows between them. By disregarding countries that do not trade with each other, these studies give up important information contained in the data, and produce biased estimates as a result, since the zeros are generally not randomly distributed. This is why truncating the sample should be avoided as a matter of practice (Westerlund and Wilhelmsson 2011).

Other empirical studies do not delete zero trade flows, but modify the dependent variable using $\ln(Y_{ij}+1)$ or $\ln(Y_{ij}+0.1)$ to accommodate the log transformation. These methods will yield generally inconsistent estimates where the severity of these inconsistencies will depend on the model and the specific characteristics of the sample used (Santos Silva and Tenreyro 2006). Hence it is not advisable to estimate the unknown parameters from the log-linear gravity model.

A natural solution to these problems is to estimate the gravity model directly from its multiplicative form. Since this removes the need to linearise the model by using logarithms, the problem with zero trade observations disappears. In doing so, note that the multiplicative gravity relationship can be written as the exponential function $\exp[\ln \beta_0 + \beta_1 \ln X_i + \beta_2 \ln X_j + \beta_3 \ln D_{ij}]$, interpreted as the conditional expectation of Y_{ij} given X_i, X_j and D_{ij}, denoted $\mathrm{E}\left[Y_{ij}|X_i, X_j, D_{ij}\right]$, as shown in Eq. (5):

$$\mu_{ij} = \mathrm{E}\left[Y_{ij}|X_i, X_j, D_{ij}\right] = \exp\left[\ln \beta_0 + \beta_1 \ln X_i + \beta_2 \ln X_j + \beta_3 \ln D_{ij}\right]. \quad (5)$$

The advantage of this specification is that the coefficients β_1, β_2 and β_3 on the logged variables X_i, X_j and D_{ij} — in the exponential relationship involving non-logged flow magnitudes as the dependent variable (Y_{ij}) — can be interpreted as the elasticity of the conditional expectation of Y_{ij} with respect to X_i, X_j and D_{ij}.

For convenience, Eq. (5) may be written in short-form as

$$\mu_k = \mathrm{E}[y_k|\boldsymbol{z}_k] = \exp(\boldsymbol{z}_k\boldsymbol{\beta}) \qquad k = 1, ..., N \tag{6}$$

where y_k denotes the k-th element of the N-by-1 vector of trade flows for the origin-destination pairs of countries. The conditional mean μ_k depends on covariates $\boldsymbol{z}_k$ with associated parameter vector $\boldsymbol{\beta}$.

3 The Poisson gravity model specification

One way to estimate the multiplicative gravity equation is based on the Poisson probability model specification, with the probability density given by

$$\mathrm{Prob}\left[y_k|\boldsymbol{z}_k\right] = \frac{\exp\left(-\mu_k\right)\mu_k^{y_k}}{y_k!} \tag{7}$$

where μ_k is specified as

$$\mu_k = \exp\left(\boldsymbol{z}_k\boldsymbol{\beta}\right). \tag{8}$$

The model has the convenient property that

$$\mathrm{E}\left[y_k|\boldsymbol{z}_k\right] = \mu_k. \tag{9}$$

Given independent observations the vector of parameters of interest, $\boldsymbol{\beta}$, can be estimated by maximizing the log likelihood function

$$\mathcal{L}(\boldsymbol{\beta}) = \sum_{k=1}^{N}\left[-\exp(\boldsymbol{z}_k\boldsymbol{\beta}) + (\boldsymbol{z}_k\boldsymbol{\beta})y_k - \ln(y_k!)\right]. \tag{10}$$

The Poisson maximum likelihood estimator is the solution to the first-order condition

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\beta}} = \sum_{k=1}^{N}\left[y_k - \exp(\boldsymbol{z}_k\boldsymbol{\beta})\right]\boldsymbol{z}_k = 0. \tag{11}$$

The essential requirement for the consistency of the maximum likelihood estimator is correct specification of the conditional mean. This is satisfied if $\mathrm{E}\left[y_k|\boldsymbol{z}_k\right] = \exp(\boldsymbol{z}_k\boldsymbol{\beta})$. The data do not have to follow the Poisson distribution and the dependent variable does not have to be integer for the estimator based on the Poisson likelihood equation (10) to be consistent (see Gourieroux et al. 1984).

Given this robustness to distributional assumptions, one can still use Eq. (11), even if the data generating process for y_k is not the Poisson. If an alternative data generating process is employed, the estimator defined by the Poisson likelihood equation is termed the Poisson pseudo maximum likelihood estimator. This terminology means that the PPML estimator is like the Poisson maximum likelihood estimator in that the Poisson model is taken to motivate the first-order condition defining the estimator, but is unlike so far that the data generating process used to obtain the distribution of the estimator does not need to be the Poisson (Cameron and Trivedi 1998, pp. 63-64). Gourieroux et al. (1984) demonstrate how pseudo maximum likelihood estimates of parametric models having finite variances will generally be consistent as long as the first-order conditional moments (that is, conditional means) are correctly specified. Even if a log likelihood function is per se misspecified, as long as its corresponding score equations have zero expectation under the true data generating process, the resulting parameter estimates will be consistent and asymptotically normal (Manning and Mullahy 2001).

The Poisson pseudo maximum likelihood estimator also has the advantage of being very well behaved, since the Hessian matrix

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = -\sum_{k=1}^{N} \exp\left(\boldsymbol{z}_k \boldsymbol{\beta}\right) \boldsymbol{z}_k' \boldsymbol{z}_k \tag{12}$$

is negative definite for all $\boldsymbol{z}$ and $\boldsymbol{\beta}$. This facilitates estimation and guarantees the uniqueness of the maximum, if it exists (see Gourieroux et al. 1984). Hence, parametric estimation is relatively simple and quasi-Newton algorithms generally converge fast, even for relatively large N.

An important implicit assumption of the Poisson gravity model is the equality between the conditional mean and the conditional variance, that is: $\mathrm{E}\left[y_k|\boldsymbol{z}_k\right] = \mathrm{var}\left[y_k|\boldsymbol{z}_k\right]$. If this assumption does not hold, then the ML coefficient estimates are consistent, but not efficient. The standard errors will be biased downward, and the inference should be based on a robust covariance matrix estimator (see Gourieroux et al. 1984 for details).

An alternative, however, is to specify the variance in a more accurate way. The negative binomial gravity model provides an obvious model specification (see, for example Fischer et al. 2006) to handle the extra variance. This probability distribution can be written as

$$\text{Prob}\left[y_k|\boldsymbol{z}_k\right] = \frac{\Gamma\left(y_k + \alpha^{-1}\right)}{\Gamma(y_k+1)\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_k}\right)^{1/\alpha} \left(\frac{\mu_k}{\alpha^{-1}+\mu_k}\right)^{y_k} \tag{13}$$

where $\Gamma(\cdot)$ is the gamma function, and α is an ancillary parameter indicating the degree of overdispersion to be estimated along with $\boldsymbol{\beta}$. The larger α is, the larger is the degree of overdispersion. The model converges to a Poisson gravity model if α is close to zero. A test of the Poisson distribution may be carried out by testing the hypothesis $\alpha = 0$ using the Wald or likelihood ratio test (see Greene 1997).

The negative binomial can be derived from a Poisson distribution in which the μ_k are distributed as a gamma random variable (Gourieroux et al. 1984, Greene 1997). The first two moments of the negative binomial distribution are given by

$$\mathrm{E}\left[y_k|\boldsymbol{z}_k\right] = \mu_k = \exp(\boldsymbol{z}_k\boldsymbol{\beta}) \tag{14}$$

$$\mathrm{var}\left[y_k|\boldsymbol{z}_k\right] = \mu_k(1+\alpha\mu_k) = \exp(\boldsymbol{z}_k\boldsymbol{\beta})\left(1+\alpha\exp(\boldsymbol{z}_k\boldsymbol{\beta})\right) \tag{15}$$

so that the expected value of the observed trade in the negative binomial gravity model is the same as in the Poisson gravity model, but the variance is specified as a function of both the conditional mean and the dispersion parameter α, incorporating unobserved heterogeneity into the conditional mean (Long 1997). Since μ and $\boldsymbol{\beta}$ are positive, $\mathrm{var}\left[y_k|\boldsymbol{z}_k\right]$ is greater than $\mathrm{E}\left[y_k|\boldsymbol{z}_k\right]$.

The negative binomial distribution belongs to the family of linear exponential distributions. Hence, the first-order condition for the NBPML estimator (see Gourieroux et al. 1984) is

$$\sum_{k=1}^{N} \boldsymbol{z}_k \frac{y_k - \exp(\boldsymbol{z}_k\boldsymbol{\beta})}{1+\alpha\exp(\boldsymbol{z}_k\boldsymbol{\beta})} = 0. \tag{16}$$

But as noticed by Bosquet and Boulhol (2010) the estimator is inappropriate when applied to a continuous dependent variable, such as trade flows measured in terms of import or export trade flows, for which the choice of the unit measure is arbitrary.

4 A zero-inflated extension of the Poisson gravity model

In recent years, it has become increasingly recognized that the level of trade between countries is frequently zero. Small countries may not have trade relations with all possible trading partners or because statistical offices do not report trade flows below a certain threshold. Non-randomly missing trade flows require the estimation of a zero-inflated gravity model that allows for two types of zero trade flows: type i zeros (sometimes called structural) and type ii zeros (sometimes called incidental).

The zero-inflated Poisson (ZIP) gravity model provides one way to model excess zero flows. Martin and Pham (2008), and Burger et al. (2009)) have proposed the zero-inflated extension of the Poisson gravity model for situations where the data generating process results into too many zeros. The model may be viewed as a "two-part" extension, in which the distribution of the outcome is approximated by mixing two component distributions (Muthén and Shedden 1999). The zero-inflated part of the model consists of a qualitative-dependent model to determine the probability of whether a particular origin-destination trade flow will be zero or positive. The second part contains the standard Poisson gravity model to estimate the relationship between trade flows and explanatory variables for each trade flow that has a non-zero probability (Leung and Yu 1996). More formally, the ZIP gravity model has the following probability function

$$\text{Prob}\,[y_k|\boldsymbol{z}_k, \boldsymbol{x}_k] = \begin{cases} \theta_k(\boldsymbol{x}_k) + [1-\theta_k(\boldsymbol{x}_k)]\exp(-\mu_k) & y_k = 0 \\ [1-\theta_k(\boldsymbol{x}_k)]\,\frac{\exp(-\mu_k)\mu_k^{y_k}}{y_k!} & y_k > 0 \end{cases} \tag{17}$$

where $\boldsymbol{x}_k$ is a vector of covariates defining the probability θ_k of extra zeros, $\theta_k \in [0,1]$. Following Lambert (1992), we can model $\theta_k(\boldsymbol{x}_k)$ using a logit model given by

$$\theta_k(\boldsymbol{x}_k) = \frac{\exp(\boldsymbol{x}_k\boldsymbol{\gamma})}{1+\exp(\boldsymbol{x}_k\boldsymbol{\gamma})}. \tag{18}$$

Although the logistic functional form is convenient, generalizations of the logistic functional form can be used as well (Cameron and Trivedi 1998). Note that the $\boldsymbol{x}$'s may be the same as the $\boldsymbol{z}$'s, but the covariates can have different effects on the $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ parameters (Long 1997).

The log likelihood function of the ZIP gravity model for a sample of N independent observation tuples $(y_k, \boldsymbol{x}_k, \boldsymbol{z}_k)$ is

$$\begin{aligned}\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = & \sum_{k=1}^{N} d_k \ln\left[\exp(\boldsymbol{x}_k\boldsymbol{\gamma}) + \exp(-\exp(\boldsymbol{z}_k\boldsymbol{\beta}))\right] \\ & + \sum_{k=1}^{N} (1 - d_k)\left[y_k \boldsymbol{z}_k\boldsymbol{\beta} - \exp(\boldsymbol{z}_k\boldsymbol{\beta}) - \ln(y_k!)\right] \\ & - \sum_{k=1}^{N} \ln\left[1 + \exp(\boldsymbol{x}_k\boldsymbol{\gamma})\right]\end{aligned} \tag{19}$$

where d_k denotes an indicator function defined by

$$d_k = \begin{cases} 1 & \text{if } y_k = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Because the model has a finite mixture structure, the maximization of the log likelihood function can use the EM algorithm as suggested by Lambert (1992), but direct estimation by means of iterative optimization methods such as Gauss-Newton, Newton-Raphson or other numerical methods is also possible. If the zero-inflated Poisson gravity model is correctly specified, maximum likelihood theory guarantees that these estimates are consistent and asymptotically efficient if they exist (Cameron and Trivedi 1998).

The first-order conditions for the Poisson pseudo maximum likelihood estimator (Staub and Winkelmann 2013) are

$$\sum_{k=1}^{N} \left[y_k - \frac{\exp(\boldsymbol{z}_k\boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}_k\boldsymbol{\gamma})}\right] \boldsymbol{z}_k = 0 \tag{21}$$

$$-\sum_{k=1}^{N} \left[y_k - \frac{\exp(\boldsymbol{z}_k\boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}_k\boldsymbol{\gamma})}\right] \frac{\exp(\boldsymbol{x}_k\boldsymbol{\gamma})}{1 + \exp(\boldsymbol{x}_k\boldsymbol{\gamma})} \boldsymbol{x}_k = 0. \tag{22}$$

This zero-inflated PPML estimator is consistent even if the true data generating process is not a Poisson distribution, as is by definition the case with excess zero trade flows. The gain of pseudo maximum likelihood estimation relative to the full maximum likelihood estimation of the zero-inflated Poisson gravity model is robustness to misspecification while the main cost is a loss of precision.

5 An eigenvector spatial filtering variant of the Poisson gravity model

As already noted in the introductory section, estimating the parameters of the Poisson gravity model (without or with zero-inflation) by pseudo maximum likelihood methods would only be justified statistically if we believed that the trade flows were independent observations. Such an assumption, however, is generally not valid, since flows are fundamentally spatial in nature. Spatial dependence is more likely than spatial independence. One way to relax this independence assumption is by incorporating spatial dependence into the Poisson version of the gravity model. Another is eigenvector spatial filtering, originally developed for area-based data (Griffith 2003) and later extended to flow data (see Chun 2008, Fischer and Griffith 2008, and Chun and Griffith 2011).

Eigenvector spatial filtering relies on the spectral decomposition of a N-by-N spatial weight matrix W into eigenvalues and eigenvectors, and then uses a subset of the eigenvectors as additional explanatory variables in the Poisson gravity model. To capture spatial dependence between origin-destination trade flows from countries neighbouring both the origin and destination locations of trade (labelled origin-destination dependence by LeSage and Pace 2008), we define the N-by-N spatial weight matrix as

$$W = W_n \otimes W_n. \tag{23}$$

W_n is a doubly stochastic n-by-n spatial weight matrix that describes the spatial neighbourhood relationships between the n countries. Neighbours may be defined using contiguity or measures of spatial proximity such as cardinal distance or ordinal distance. The spatial weight matrix has – by convention – zeros in the main diagonal, and non-negative elements in the off-diagonal cells. $\otimes$ denotes the Kronecker product.

In this study we follow Pace et al. (2013) to specify W_n using m individual n-by-n neighbour matrices M_t $(t = 1, ..., m)$ with each M_t containing a one in each row if an observation is the t-th nearest neighbour to that observation, and zero otherwise. We use a geometric decay parameter δ to allow for declining influence of higher order neighbour matrices in the overall n-by-n spatial weight matrix W_n. Thus

$$W_n = D(M + M')D \tag{24}$$

$$M = M_1(1-\delta) + M_2(1-\delta)^2 + ... + M_m(1-\delta)^m \tag{25}$$

where the diagonal scaling matrix D is calculated iteratively in order to make W_n doubly stochastic (Pace and LeSage 2002). Note that $M + M'$ is by construction symmetric. In accordance with Pace et al. (2013) we set $\delta = 0.05$ and $m = 30$. This gives the 15th nearest neighbour roughly half the influence of the first nearest neighbour.

Spatial filtering uses the spectral decomposition of the spatial weight matrix W

$$W = E\Lambda E' \tag{26}$$

where E is an N-by-N matrix of eigenvectors, and Λ is an N-by-N matrix containing the corresponding eigenvalues on the diagonal. The orthogonality characteristics of eigenvectors make the spectral decomposition useful for lower rank approximations to W (see Pace et al. 2013). The idea is to keep all the eigenvectors associated with the largest magnitude eigenvalues and to discard the rest. This involves partitioning the eigenvalues and eigenvectors into two sets, a set of eigenvectors associated with the largest Q eigenvalues and a set of eigenvectors associated with the smallest $N-Q$ eigenvalues of W. To identify and optimise the subset of Q eigenvectors we follow Tiefelsdorf and Griffith (2007) using a stepwise Poisson regression selection technique. The Q eigenvectors are then used as additional explanatory variables in Eq. (5) so that

$$\mu_{ij} = \exp\left[\ln\beta_0 + \beta_1 \ln X_i + \beta_2 \ln X_j + \beta_3 \ln D_{ij} + \sum_{q=1}^{Q} E_q \Phi_q\right] \tag{27}$$

where E_q denotes the q-th eigenvector and Φ_q its associated regression coefficient. The term

$$\exp\left[\sum_{q=1}^{Q} E_q \Phi_q\right] \tag{28}$$

is called a spatial filter that accounts for origin-destination dependence with a linear combination of eigenvectors. Filtering out spatial aspects in the flow data reduces potential bias in parameter estimates.

6 An empirical application

In this section, we use the eigenvector spatial filtering variants of the Poisson gravity model (without and with zero-inflation) along with pseudo maximum likelihood methods, to quantitatively assess the determinants of international trade flows, uncovering significant differences in the role of various distance measures from those predicted by the standard Poisson gravity model specification. For the sake of completeness we also compare the results obtained by the Poisson gravity model specification with those produced by the log-normal model, and, moreover, illustrate that the NBPML estimator is inappropriate when applied to trade flows.

The data

Our data set consists of 21,170 observations on bilateral flows (146 x 145 country pairs). The countries are listed in Table A in the appendix. Instead of constructing symmetric trade flows by combining exports and imports for each country pair, we use unidirectional export trade values, measured in terms of millions of US dollars, as an indicator of the trade volume between countries in the year 2000. Information on bilateral exports comes from Feenstra et al. (2005). The data are in multiple dyadic format. That is, each country pair appears twice corresponding to the export of one country to another and vice versa. About 53% of all observations are zero trade flows. Most of the zero trade flows reflect a true absence of trade, rather than non-repeating, omissions or random errors. Helpman et al. (2008) attribute these zeros to failure to meet the fixed costs associated with establishing trade flows.

Data on real *gdp* comes from World Bank's (2002) *World Development Indicators*. We distinguish four types of barriers or enhancements typically controlled for in gravity studies of trade patterns: geographical distance and language barriers, common land border and trade policy enhancement, measured in terms of preferential trade agreements. Bilateral distance between countries is computed based on the great circle distance between their capital cities. Data on dummies indicating a common land border and common language (official and second languages) are constructed from CIA's (2011) *World Factbook*. The data on language ties are presented in Table B of the appendix. Information on preferential trade agreements comes from WTO's (2014) regional trade agreement database. The list of preferential trade agreements and stronger forms of trade agreements considered in the analysis is summarized in Table C in the appendix. Finally, Table D in the appendix lists descriptive statistics for the covariates in the gravity model specifications.

Table 1 about here

Results

Table 1 presents the estimation outcomes of the gravity model, using OLS and PPML estimation. The first column reports OLS estimates using the logarithm of exports as the dependent variable. As noted in section 2, this regression excludes pairs of countries with zero bilateral trade (only 47% of the sample have positive export flows). The second column presents PPML estimates using only the subsample of positive trade pairs, while the third column shows the Poisson results for the whole sample (i.e. including zero trade pairs).

The first point to note is that the PPML estimated coefficients are remarkably similar using the whole sample and using the positive trade subsample. All the coefficients differ significantly from those resulting from using OLS. This indicates that heteroscedasticity is responsible for the differences between the PPML results and those of OLS using merely the observations with positive exports. Further evidence on the importance of heteroscedasticity is indicated by the two-degrees-of-freedom special case of White's test for heteroscedasticity (see Wooldridge 2001) that yields a test statistic of 0.022 and a p-value less than 0.01. This implies that the null hypothesis of homoscedastic errors is rejected.

The PPML estimates reveal that the coefficients on importer's and exporter's gross domestic product in the gravity model are not, as generally believed, close to one. The estimated *gdp* elasticities are around 0.8 (s.e. $\approx$ 0.03). OLS generates significantly larger estimates (exporter's *gdp*: 0.97, s.e. = 0.01; importer's *gdp*: 0.93, s.e. = 0.01). The role of geographical distance as trade deterrent is significantly larger under OLS. The estimated elasticity is -0.77 (s.e. = 0.02), whereas the PPML estimate is -0.49 (s.e.=0.06). This lower estimate suggests a smaller role for transport costs in the determination of trade pattern. The Poisson estimates, moreover, indicate that – after controlling for bilateral distance – sharing a border generates a substantial effect. The Poisson gravity model specification predicts that trade between two contiguous countries is 90,9% [= $(e^{0.647} - 1)100$] larger than trade between countries that do not share a border (OLS: 127.8%). Free trade agreements play a smaller – but still substantial – role, according to the log-linear gravity model. The OLS estimate suggests that trade agreements raise expected bilateral trade by 34.9%, while PPML estimates indicate that trade agreements have no significant effect on trade. We also find significant discrepancies in the role of common language. The Poisson pseudo maximum likelihood method estimates a smaller effect on trade, approximately one third of that indicated by OLS.

Table 2 about here

Some researchers consider other PML estimators based on non-Poisson distributions such as the negative binomial (see, for example, Burger et al. 2009). The NBPML estimator is appealing. But Table 2 provides clear evidence that the NBPML estimated parameters artificially depend on whether the trade flows are measured in thousands, millions, billions or trillions of USD. This scale dependence of the NBPML estimator invalidates the choice of the negative binomial specification of the gravity model. It is worth noting that when the flows are measured in small units, such as thousands of USD, the NBPML estimator tends to converge to the PPML estimator.

Table 3 about here

We now turn to the spatial filtering variant of the Poisson gravity model that filters out origin-destination dependence between flows to reduce potential bias in the parameter estimates. Table 3 reports the PPML estimates reflecting results obtained by using $m = 30$ and $\delta = 0.05$ to specify the spatial weight matrix W_n between the countries, in comparison to those of the unfiltered Poisson gravity model version. Table E in the appendix illustrates the robustness of the results. Performance of the model is expressed in terms of conventional statistical measures of goodness of fit, such as the log likelihood divided by N, and R^* defined as the correlation between the fitted and observed values of the dependent variable.

The first three columns of the table present the results that account for origin-destination dependence in the flows, and the final three report those that do not. The parameter estimates are given in the first and fourth column, followed by the standard errors in the second and fifth, and the p-values in the third and sixth columns. Comparing the results of the two model specifications, we observe some substantial differences. First, trading countries having a common language exhibit a positive effect on trade, but the effect is exaggerated by roughly one third when neglecting spatial dependence. Second, the unfiltered Poisson gravity model indicates that preferential trade agreements play no significant role, while estimates resulting from the spatial filtering model specification suggest that preferential trade agreements raise expected bilateral trade by 29.8%. These differences in the estimates clearly indicate that the biases generated by neglecting spatial dependence can be substantial, yielding misleading inferences and, perhaps, erroneous policy decisions.

Table 4 about here

The PPML parameter estimates from the zero-inflated extension of the spatially filtered Poisson gravity model are presented in Table 4. While we do not discuss the interpretation of individual parameters, there are several things to note. The results clearly show that the very same variables that impact export volumes from country i to country j also impact the probability that i exports to j. The signs of the $\boldsymbol{\gamma}$s in the formation of bilateral trade relationships are the opposite of the $\boldsymbol{\beta}$s, except for the common land border variable. The binary process is predicting membership in the group of country pairs that must have a zero flow. A positive $\boldsymbol{\gamma}$ is associated with higher probability of not trading at all. If variables that positively affect the expected flows (i.e., those with positive $\boldsymbol{\beta}$s) also positively affect the chances of being in the group where positive export flows are possible, then the $\boldsymbol{\beta}$s and $\boldsymbol{\gamma}$s would have opposite signs. If both the $\boldsymbol{\gamma}$s and $\boldsymbol{\beta}$s are positive, this can be interpreted as the variable affecting the chance not to trade at all positively, but once, despite this influence, trade is established, then having a positive influence on the volume of trade. This is the case with the effect of a common land border. It raises the volume of trade, but reduces the probability of trading. This finding is in accordance with Helpman et al. (2008), and may be attributed to the effect of territorial border conflicts that suppress trade between neighbours. In the absence of such conflicts, common land borders enhance trade.

The PPML $\boldsymbol{\beta}$-parameter estimates from the spatial filtering Poisson gravity model and its zero-inflation variant are remarkably similar. The only exception is the parameter estimate of the common land border variable. The important question is whether the spatially filtered ZIP model provides any improvement over its counterpart model without zero-inflation. The log likelihood function value is higher, but since the models are not nested, the log likelihood values are not directly comparable. Greene (1994) proposes for this case to use the test statistics developed by Vuong (1989) for non-nested models. Without going into mathematical details, the test compares the predicted probabilities of trade flows for two different models, and the resulting V statistic is asymptotically normally distributed. Based on the Vuong test, we find that the spatial filtering Poisson gravity model with zero-inflation is clearly preferred to its counterpart specification without zero-inflation ($V = 3.371$, $p < 0.001$).

7 Closing remarks

In this paper, we argue that the standard practice used to log-linearise the gravity model and estimate the parameters of interest by least squares is inappropriate. The basic problem with this approach is that in the presence of heteroscedasticity log-linearisation leads to inconsistent estimates. An additional problem is that log-linearisation is incompatible with the existence of zero flows in the data, and disregarding zero trade flows produces biased results, since the zeros are generally not randomly distributed.

In recent years the Poisson probability specification of the gravity model along with pseudo maximum likelihood methods has received increasing popularity to address these estimation problems. This approach is robust to different patterns of heteroscedasticity, and, in addition, provides a natural way to deal with the zero problem in trade flow data. We argue that estimating the model parameters by means of PPML leads to consistent, but biased parameter estimates if spatial dependence between origin-destination flows is ignored.

To overcome this estimation problem we suggest eigenvector spatial filtering variants of the Poisson gravity model (without or with zero-inflation) along with pseudo maximum likelihood estimation. We use this approach to re-estimate the gravity equation, and find strong evidence that Poisson pseudo maximum likelihood estimation based on the unfiltered Poisson gravity model produces biased parameter estimates. Differences indicate that the biases generated by neglecting origin-destination dependence between trade flows can be substantial, yielding misleading inferences and, perhaps, erroneous policy recommendations. Finally, it is worth noting that – based on the Vuong test – the zero-inflated extension of the spatial filtering variant that provides a way to model excess zero flows is found to be clearly preferable to its counterpart spatially filtered specification without zero-inflation.

References

Anderson JE (1979) A theoretical foundation for the gravity equation. *American Economic Review* 69(1), 106–116

Anderson JE and van Wincoop E (2003) Gravity with gravitas: A solution to the border puzzle. *American Economic Review* 93(1), 170–192

Bosquet C and Boulhol H (2010) Scale-dependence of the negative binomial pseudo-maximum likelihood estimator. *Documents de Travail du Centre d'Economie de la Sorbonne, 2010.90*

Broyden CG (1970) The convergence of a class of double-rank minimization algorithms 2. The new algorithm. *IMA Journal of Applied Mathematics* 6(3), 222–231

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

Cameron AC and Trivedi PK (1998) *Regression analysis of count data*. Cambridge University Press, Cambridge

Chun Y (2008) Modeling network autocorrelation within migration flows by eigenvector spatial filtering. *Journal of Geographical Systems* 10(4), 317–344

Chun Y and Griffith DA (2011) Modeling network autocorrelation in space-time migration flow data: An eigenvector spatial filtering approach. *Annals of the Association of American Geographers* 101(3), 523–536

CIA (2011) *CIA World factbook*. CIA, Washington D.C.

Feenstra RC, Lipsey RE, Deng H, Ma AC and Mo H (2005) World trade flows: 1962-2000. Technical Report, National Bureau of Economic Research, NBER Working Paper No. 11040

Fischer MM and Griffith DA (2008) Modeling spatial autocorrelation in spatial interaction data: An application to patent citation data in the European Union. *Journal of Regional Science* 48(5), 969–989

Fischer MM, Scherngell T and Jansenberger E (2006) The geography of knowledge spillovers between high-technology firms in Europe: Evidence from a spatial interaction modeling perspective. *Geographical Analysis* 38(9), 288–309

Fletcher R (1970) A new approach to variable metric algorithms. *The Computer Journal* 13(3), 317–322

Frankel JA (1997) Regional trading blocs in the world economic system. Institute for International Economics, Washington DC

Goldfarb D (1970) A family of variable-metric methods derived by variational means. *Mathematics of Computation* 24(109), 23–26

Gourieroux C, Monfort A and Trognon A (1984) Pseudo-maximum likelihood methods: Applications to Poisson models. *Econometrica* 52(3), 701–720

Greene WH (1994) Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. Working Paper 94-10, Stern School of Business, New York University

Greene WH (1997) *Econometric analysis*. 7th edition, Pearson Education, Essex

Griffith DA (2003) *Spatial autocorrelation and spatial filtering: Gaining understanding through theory and scientific visualization*. Springer, Berlin Heidelberg New York

Griffith DA and Fischer MM (2013) Constrained variants of the gravity model and spatial dependencies: Model specification and estimation issues. *Journal of Geographical Systems* 13(3), 1–27

Helpman E, Melitz M and Rubinstein Y (2008) Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics* 123(2), 441–487

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

LeSage JP and Pace RK (2008) Spatial econometric modeling of origin-destination flows. *Journal of Regional Science* 48(5), 941–967

Leung SF and Yu S (1996) On the choice between sample selection and two-part models. *Journal of Econometrics* 72(1), 197–229

Lin G and Zhang T (2007) Loglinear residual tests of Moran's I autocorrelation and their applications to Kentucky breast cancer data. *Geographical Analysis* 39(3), 293–310

Linders GM, Burger MJ and van Oort FG (2008) A rather empty world: The many faces of distance and the persistent resistance to international trade. *Cambridge Journal of Regions, Economy and Society* 1(3), 439–458

Liu X (2009) GATT/WTO promotes trade strongly: Sample selection and model specification. *Review of International Economics* 17(3), 428–446

Long JS (1997) *Regression models for categorical and limited dependent variables.* Sage, Thousand Oaks, London and New Delphi

Manning WG and Mullahy J (2001) Estimating log models: To transform or not to transform? *Journal of Health Economics* 20(4), 461–494

Martin W and Pham C (2008) Estimating the gravity model when zero trade flows are important. *World Bank. Mimeo*

Martínez-Zarzoso I (2013) The log of gravity revisited. *Applied Economics* 45(3), 311–327

Muthén B and Shedden K (1999) Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 55(2), 463–469

Pace RK and LeSage JP (2002) Semiparametric maximum likelihood estimates of spatial dependence. *Geographical Analysis* 34(1), 76–90

Pace RK, LeSage JP and Zhu S (2013) Interpretation and computation of estimates from regression models using spatial filtering. *Spatial Economic Analysis* 8(3), 352–369

Santos Silva JMC and Tenreyro S (2006) The log of gravity. *The Review of Economics and Statistics* 88(4), 641–658

Santos Silva JMC and Tenreyro S (2010) On the existence of the maximum likelihood estimates in Poisson regression. *Economics Letters* 107(2), 310–312

Santos Silva JMC and Tenreyro S (2011) Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters* 112(2), 220–222

Shanno DF (1970) Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation* 24(111), 647–656

Staub KE and Winkelmann R (2013) Consistent estimation of zero-inflated count models. *Health Economics* 22(6), 673–686

Tiefelsdorf M and Griffith DA (2007) Semiparametric filtering of spatial autocorrelation: The eigenvector approach. *Environment and Planning A* 39(5), 1193–1221

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

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2), 307–333

Westerlund J and Wilhelmsson F (2011) Estimating the gravity model without gravity using panel data. *Applied Economics* 43(6), 641–649

Wooldridge J (2001) *Econometric analysis of cross section and panel data*. MIT Press, Cambridge

World Bank (2002) *World development indicators 2002*. World Bank Publications, Oxford University Press, New York

WTO (2014) *WTO regional trade agreements database*. WTO, accessed on: 2014-07-01

Table 1: Parameter estimates: OLS versus PPML estimation

Estimator	OLS	PPML	
Dependent variable	$\ln(Y_{ij} > 0)$	$Y_{ij} > 0$	Y_{ij}
Log exporter's *gdp*	0.969 ***	0.798 ***	0.817 ***
	(0.009)	(0.005)	(0.022)
Log importer's *gdp*	0.926 ***	0.829 ***	0.852 ***
	(0.008)	(0.006)	(0.025)
Log distance [in 1,000 km]	-0.771 ***	-0.467 ***	-0.487 ***
	(0.019)	(0.016)	(0.061)
Common land border	0.823 ***	0.656 ***	0.647 ***
	(0.100)	(0.108)	(0.179)
Free trade agreement	0.300 ***	0.047	0.059
	(0.047)	(0.037)	(0.128)
Common language	0.911 ***	0.460 ***	0.415 ***
	(0.047)	(0.117)	(0.116)
Constant	-43.538 ***	-35.806 ***	-36.993 ***
	(0.352)	(1.019)	(0.962)
R^*	0.831	0.852	0.853
Number of observations	9,957	9,957	21,170

Notes: Standard errors in parentheses represent robust White standard errors. R^* is defined as the overall fit of the model in terms of the correlation between the fitted and observed values of the dependent variable. *** denotes statistical significance at the one percent level. The first-order condition for the PPML estimator has been solved using the optim() function in R, with the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno algorithm (see Broyden 1970, Fletcher 1970, Goldfarb 1970 and Shanno 1970).

Table 2: Scale dependence of the NBPML estimator

	Trade flows measured in terms of			
	thousands USD	millions USD	billions USD	trillions USD
Log exporter's *gdp*	0.792 ***	1.083 ***	0.969 ***	1.071 ***
	(0.193)	(0.047)	(0.018)	(0.044)
Log importer's *gdp*	0.825 ***	0.964 ***	1.024 ***	1.144 ***
	(0.072)	(0.109)	(0.020)	(0.045)
Log distance [in 1,000 km]	-0.488 ***	-0.551 ***	-0.763 ***	-0.747 ***
	(0.150)	(0.082)	(0.048)	(0.086)
Common land border	0.646 **	0.668 ***	0.498 ***	0.760 ***
	(0.286)	(0.196)	(0.198)	(0.254)
Free trade agreement	0.063	0.132	0.126	-0.227
	(0.184)	(0.115)	(0.090)	(0.222)
Common language	0.415	1.066 ***	0.955 ***	0.706 ***
	(0.821)	(0.194)	(0.116)	(0.185)
Constant	-30.080 ***	-46.804 ***	-52.403 ***	-61.355 ***
	(3.188)	(1.730)	(0.819)	(1.808)
Overdispersion [α]	83.451 ***	4.437 ***	0.884	0.269
	(0.495)	(0.488)	(0.236)	(0.000)
Log likelihood/N	-1.61E+02	-9.37E+01	-1.89E+01	-0.03E+01
R^*	0.728	0.779	0.754	0.779
Number of observations	21,170	21,170	21,170	21,170

Notes: Standard errors in parentheses represent robust White standard errors. R^* is defined as the overall fit of the model in terms of the correlation between the fitted and observed values of the dependent variable. *** and ** denote statistical significance at the one percent level and five percent level, respectively. The first-order condition for the NBPML estimator has been solved using the optim() function in R, with the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno algorithm (see Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970).

Table 3: Parameter estimates: The spatially filtered Poisson gravity model compared with its unfiltered counterpart model

	Spatially filtered Poisson gravity model			Unfiltered Poisson gravity model		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Log exporter's *gdp*	0.758	0.019	0.000	0.817	0.022	0.000
Log importer's *gdp*	0.807	0.022	0.000	0.852	0.025	0.000
Log distance [in 1,000 km]	-0.519	0.056	0.000	-0.487	0.061	0.000
Common land border	0.626	0.138	0.000	0.647	0.179	0.000
Free trade agreement	0.261	0.114	0.011	0.059	0.128	0.323
Common language	0.287	0.098	0.002	0.415	0.116	0.000
Constant	-24.190	0.829	0.000	-36.993	0.962	0.000
Adjusted Moran's I	0.078	0.034	0.082	0.343	0.001	0.000
Log likelihood/N			-9.30E+01			-1.37E+05
R^*			0.900			0.853
Number of eigenvectors			52			—
Number of observations			21,170			21,170

Notes: Standard errors represent robust White standard errors. Spatial filtering is based on the N by N spatial weight matrix $W = W_n \otimes W_n$, where W_n is specified using $m = 30$ individual n-by-n neighbour matrices and the geometric decay parameter $\delta = 0.05$. Following Tiefelsdorf and Griffith (2007) we identified $Q = 52$ eigenvectors. An adjusted version of Moran's I statistic is used (see Lin and Zhang 2007). R^* is defined as the overall fit of the model in terms of the correlation between the fitted and observed values of the dependent variable. The first-order conditions for the estimators have been solved using the optim() function in R, with the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno algorithm (see Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970).

Table 4: Parameter estimates: The spatially filtered Poisson gravity model with zero-inflation

	The spatially filtered Poisson gravity model with zero-inflation					
	Logit			Poisson		
	Coefficient	Std. error	p-value	Coefficient	Std.error	p-value
Log exporter's *gdp*	-0.950	0.330	0.002	0.763	0.016	0.000
Log importer's *gdp*	-0.917	0.369	0.006	0.804	0.016	0.000
Log distance [in 1,000 km]	0.102	0.016	0.000	-0.539	0.028	0.000
Common land border	0.290	0.014	0.000	0.543	0.043	0.000
Free trade agreement	-0.054	0.028	0.026	0.238	0.026	0.000
Common language	-0.627	0.080	0.000	0.268	0.085	0.001
Constant	28.996	0.716	0.000	-23.997	0.862	0.000
Adjusted Moran's I				0.003	0.002	0.067
Log likelihood/N						-1.91E+06
R^*						0.852
Number of eigenvectors						52
Number of observations						21,170

Notes: Standard errors represent robust White standard errors. Spatial filtering is based on the N by N spatial weight matrix $W = W_n \otimes W_n$, where W_n is specified using $m = 30$ individual n-by-n neighbour matrices and the geometric decay parameter $\delta = 0.05$. Following Tiefelsdorf and Griffith (2007) we identified $Q = 52$ eigenvectors. An adjusted version of Moran's I statistic is used (see Lin and Zhang 2007). R^* is defined as the overall fit of the model in terms of the correlation between the fitted and observed values of the dependent variable. The first-order conditions for the estimator have been solved using the optim() function in R, with the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno algorithm (see Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970).

Appendix

Table A: List of countries

Albania	Ecuador	Latvia	Rwanda
Algeria	Egypt	Lebanon	Saudi Arabia
Angola	El Salvador	Libya	Senegal
Argentina	Equatorial Guinea	Lithuania	Seychelles
Australia	Estonia	Madagascar	Sierra Leone
Austria	Ethiopia	Malawi	Singapore
Bahamas	Fiji	Malaysia	Slovakia
Bahrain	Finland	Mali	Slovenia
Bangladesh	France-Monaco	Malta	South Africa
Barbados	Gabon	Mauritania	Spain
Belarus	Gambia	Mauritius	Sri Lanka
Belgium-Luxembourg	Germany	Mexico	St. Kitts-Nevis-Antilles
Belize	Ghana	Mongolia	Sudan
Benin	Greece	Morocco	Suriname
Bolivia	Guatemala	Mozambique	Sweden
Brazil	Guinea	Nepal	Switzerland-Liechtenstein
Bulgaria	Guinea-Bissau	Netherlands	Syria
Burkina Faso	Guyana	New Caledonia	Taiwan
Burundi	Haiti	Peru	Tanzania
Côte D'Ivorie	Italy	New Zealand	USA
Cambodia	Hong Kong	Nicaragua	Thailand
Cameroon	Honduras	Niger	Togo
Canada	Hungary	Nigeria	Trinidad and Tobago
Central African Rep.	Iceland	Norway	Tunisia
Chad	India	Oman	Turkey
Chile	Indonesia	Pakistan	Uganda
China	Iran	Panama	UK
Colombia	Iraq	Papua New Guinea	Ukraine
Congo	Ireland	Paraguay	United Arab Emirates
Costa Rica	Israel	Philippines	Uruguay
Croatia	Jamaica	Poland	Venezuela
Cyprus	Japan	Portugal	Vietnam
Czech Rep.	Jordan	Qatar	Yemen
Dem. Rep. of Congo	Kenya	Rep. of Korea	Zambia
Denmark	Kiribati	Rep. of Moldova	Zimbabwe
Djibouti	Kuwait	Romania	
Dominican Rep.	Laos	Russian Fed.	

Table B: Language ties: Common official and second languages (CIA 2011)

English	**French**	**Spanish**	**German**
Australia	Algeria	Argentina	Austria
Bahamas	Belgium-Luxembourg	Belize	Germany
Barbados	Benin	Bolivia	Switzerland-Liechtenstein
Cameroon	Burkina Faso	Chile	
Canada	Burundi	Colombia	**Arabic**
Fiji	Cameroon	Costa Rica	Algeria
Gambia	Canada	Dominican Rep.	Bahrain
Guyana	Cent. African Rep.	Ecuador	Chad
India	Chad	El Salvador	Djibouti
Ireland	Congo	Equatorial Guinea	Egypt
Jamaica	Dem. Rep. of Congo	Guatemala	Jordan
Kenya	Côte D'Ivorie	Honduras	Lebanon
Kiribati	Djibouti	Mexico	Mauritania
Malawi	Equatorial Guinea	Nicaragua	Morocco
Malta	France-Monaco	Panama	Oman
Mauritius	Gabon	Paraguay	Saudi Arabia
New Zealand	Guinea	Peru	Sudan
Nigeria	Haiti	Spain	Syria
Oman	Lebanon	Uruguay	Tanzania
Pakistan	Madagascar	Venezuela	Tunisia
Panama	Mali		United Arab Emirates
Papua New Guina	Mauritius	**Italian**	Yemen
Philippines	Morocco	Italy	Kuwait
Rwanda	New Caledonia	Libya	Iraq
Seychelles	Niger	Switzerland-Liechtenstein	Qatar
Sierra Leone	Rwanda		
Singapore	Senegal	**Slavic**	**Malay**
South Africa	Seychelles	Belarus	Indonesia
Sri Lanka	Switzerland-Liechtenstein	Bulgaria	Malaysia
St.Kitts-Nevis-Antilles	Togo	Croatia	Singapore
Suriname	Tunisia	Czech Rep.	
Tanzania		Poland	**Chinese**
Trinidad Tobago	**Portuguese**	Russian Fed.	China
Uganda	Angola	Slovenia	Hong Kong
UK	Brazil	Slovakia	Malaysia
USA	Guinea-Bissau	Ukraine	Singapore
Zambia	Mozambique		
Zimbabwe	Portugal	**Dutch**	
		Belgium-Luxembourg	
		Netherlands	
		Suriname	

Table C: Free trade and stronger forms of agreements in 2000 (WTO 2014)

APTA	**COMESA**	**GSTP**	**MERCOSUR**
Bangladesh	Burundi	Algeria	Argentina
India	Dem. Republic of Congo	Argentinia	Bolivia
Laos	Djibouti	Bangladesh	Brazil
Nepal	Egypt	Benin	Chile
Philippines	Ethiopia	Bolivia	Paraguay
Rep. of Korea	Kenya	Brazil	Uruguay
Sri Lanka	Madagascar	Cameroon	
	Mauritius	Chile	**NAFTA**
ASEAN [AFTA]	Rwanda	Colombia	Canada
Cambodia	Sudan	Ecuador	Mexico
Indonesia	Uganda	Egypt	USA
Laos	Zambia	Ghana	
Malaysia	Zimbabwe	Guinea	**PATCRA**
Philippines		Guyana	Australia
Singapore	**ECO**	India	Papua New Guinea
Thailand	Iran	Indonesia	
Vietnam	Pakistan	Iran	**SICA**
	Turkey	Iraq	Costa Rica
CAN		Libya	El Salvador
Bolivia	**ECOWAS**	Malaysia	Guatemala
Colombia	Benin	Mexico	Honduras
Ecuador	Burkina Faso	Morocco	Nicaragua
Peru	Côte D'Ivoire	Mozambique	
	Gambia	Nicaragua	**EFTA treaties**
CACM	Ghana	Nigeria	EFTA-Israel
Costa Rica	Guinea	Pakistan	EFTA-Morocco
El Salvador	Guinea-Bissau	Peru	EFTA-Turkey
Honduras	Mali	Philippines	
Nicaragua	Niger	Rep. of Korea	**EU treaties**
Guatemala	Nigeria	Singapore	EU-Iceland
	Senegal	Sri Lanka	EU-Israel
CARICOM	Sierra Leone	Sudan	EU-Norway
Bahamas	Togo	Tanzania	EU-South Africa
Barbados		Thailand	EU-Syria
Belize	**EFTA**	Trinidad and Tobago	EU-Tunisia
Dominican Rep.	Iceland	Tunisia	EU-Turkey
Guyana	Norway	Zimbabwe	
Haiti	Switzerland-Liechtenstein		**Bilateral treaties**
Jamaica		**LAIA**	Belarus-Russian Fed.
St. Kitts-Nevis-Antilles	**EU**	Argentina	Canada-Chile
Suriname	Austria	Bolivia	Canada-Israel
Trinidad and Tobago	Belgium-Luxembourg	Brazil	Chile-Mexico
	Denmark	Chile	Colombia-Mexico
CEMAC	Finland	Colombia	Fiji-Papua New Guinea
Angola	France-Monaco	Ecuador	Israel-Mexico
Burundi	Germany	Mexico	Laos-Thailand
Cameroon	Greece	Paraguay	New Zealand-Australia
Central African Rep.	Ireland	Uruguay	Turkey-Israel
Chad	Italy	Panama	Ukraine-Russian Fed.
Congo	Netherlands	Peru	
Dem. Rep. Of Congo	Portugal	Venezuela	
Equatorial Guinea	Spain		
Gabon	Sweden		
	UK		

Note: Asia Pacific Trade Agreement (APTA), Asian Free Trade Area (AFTA), Andean Community (CAN), Central American Common Market (CACM), Caribbean Community and Common Market (CARICOM), Economic Community of Central African States (CEMAC), Common Market for Eastern and Southern Africa (COMESA), Economic Cooperation Organization (ECO), Economic Community of West African States (ECOWAS), European Free Trade Agreement (EFTA), Global System of Trade Preferences among Developing Countries (GSTP), Latin American Integration Association (LAIA), North American Free Trade Agreement (NAFTA), Mercado Comun del Sur (MERCOSUR), Agreement on Trade between Australia and New Guinea (PATCRA), Central American Integration System (SICA)

Table D: Summary statistics for the covariates

	Full sample				Positive trade subsample			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Log exporter's *gdp*	23.931	2.179	18.329	29.974	25.070	2.066	18.329	29.974
Log importer's *gdp*	23.931	2.179	18.329	29.974	25.017	2.111	18.329	29.974
Log distance [in 1,000 km]	1.799	0.789	-4.949	2.992	1.736	0.825	-2.870	2.992
Common land border	0.021	0.144	0	1	0.024	0.153	0	1
Free trade agreement	0.116	0.320	0	1	0.139	0.346	0	1
Common language	0.178	0.382	0	1	0.160	0.366	0	1
Number of observations	21,170				9,957			

Table E: Robustness check: The impact of varying m and δ (used to specify W_n) on the parameter estimates

	$m = 25$			$m = 30$			$m = 35$		
	δ =0.04	δ =0.05	δ =0.06	δ =0.04	δ =0.05	δ =0.06	δ =0.04	δ =0.05	δ =0.06
Log exporter's *gdp*	0.763 ***	0.762 ***	0.761 ***	0.759 ***	0.758 ***	0.757 ***	0.756 ***	0.754 ***	0.753 ***
Log importer's *gdp*	0.803 ***	0.804 ***	0.805 ***	0.806 ***	0.807 ***	0.808 ***	0.809 ***	0.810 ***	0.811 ***
Log distance [1,000 km]	-0.501 ***	-0.506 ***	-0.510 ***	-0.515 ***	-0.519 ***	-0.524 ***	-0.528 ***	-0.533 ***	-0.537 ***
Common land border	0.542 ***	0.563 ***	0.584 ***	0.605 ***	0.626 ***	0.647 ***	0.668 ***	0.689 ***	0.710 ***
Free trade agreement	0.238	0.244	0.250	0.255	0.261	0.267	0.273	0.278	0.284
Common language	-0.267 ***	-0.272 ***	-0.277 ***	-0.282 ***	-0.287 ***	-0.292 ***	-0.297 ***	-0.302 ***	-0.307 ***
Constant	-23.996 ***	-24.045 ***	-24.093 ***	-24.142 ***	-24.190 ***	-24.239 ***	-24.287 ***	-24.336 ***	-24.384 ***

Notes: Spatial filtering is based on the N by N spatial weight matrix $W = W_n \otimes W_n$, where W_n is specified using $m = 25$, 30 and 35 individual n-by-n neighbour matrices and varying geometric decay parameters $\delta = 0.04$, 0.05 and 0.06. *** denotes statistical significance at the one percent level. The first-order conditions for the estimator have been solved using the optim() function in R, with the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno algorithm (see Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970).