Received 2 June 2022
Revised 14 June 2022
Accepted 14 June 2022

Do trade flows interact in space? Spatial origin-destination modeling of gravity

Shali Luo
Department of Economics, University of Missouri Columbia, Columbia, Missouri, USA, and
Seung-Whan Choi
Department of Political Science, University of Illinois at Chicago, Chicago, Illinois, USA

Abstract

Purpose – This study proposes spatial origin-destination threshold Tobit to address spatial interdependence among bilateral trade flows while accounting for zero trade volumes.

Design/methodology/approach – This model is designed to capture multiple forms of spatial autocorrelation embedded in "directional" trade flows. The authors apply this improved model to export flows among 32 Asian countries in 1990.

Findings – The empirical results indicate the presence of all three types of spatial dependence: exporter-based, importer-based and exporter-to-importer-based. After further considering multifaceted spatial correlation in bilateral trade flows, the authors find that the effect of conventional trade variables changes in a noticeable way.

Research limitations/implications – This finding implies that the standard gravity model may produce biased estimates if it does not take spatial dependence into account.

Originality/value – This paper attempts to offer an improved model of the standard gravity model by taking spatial dependence into account.

Keywords Trade, Spatial dependence, Tobit

Paper type Research paper

Since Tinbergen (1962) introduced a gravity equation as an empirical specification of bilateral trade flows, his model "has dominated empirical research in international trade" (Helpman *et al.*, 2008, p. 442). In its basic form, his gravity model explains trade volumes based on the economic size (often measured by real GDP) of two trading partners and the distance separating them through a functional form analogous to Newton's Law of Gravity with stochastic features. During the past decades, subsequent research has extended his basic model by including other explanatory variables to help better understand the mechanism of trade (e.g. border effect, McCallum, 1995) and evaluate institutional or policy impacts on trade flows (e.g. preferential trade agreement and membership in WTO, Feenstra, 2004). As Anderson (2011, p. 106) succinctly summarized, "[A]pplied to a wide variety of goods and factors moving over regional and national borders under differing circumstances, [the gravity model] usually produces a good fit". Given its strong explanatory power, simple formulation and easy interpretation in log-linear form, scholars view the gravity model as a successful empirical tool to examine trade flows (Anderson and van Wincoop, 2003).

This study grew out of Shali Luo's dissertation at the University of Missouri–Columbia.

International Trade, Politics and Development
Vol. 6 No. 2, 2022
pp. 46-60
Emerald Publishing Limited
e-ISSN: 2632-122X
p-ISSN: 2586-3932
DOI 10.1108/ITPD-06-2022-0009

Despite its wide popularity, several scholars have attempted to improve the gravity model over the years (e.g. Pfaffermayr, 2019; Weidner and Zylkin, 2021). In this study, we make an empirical contribution by modeling dependent variables with three nonstandard yet empirically relevant statistical features: censored, dyadic and spatially correlated data. By utilizing recent advances of spatial econometrics in modeling spatial autocorrelation for data featuring origin-destination (OD) flows (see LeSage and Pace, 2008, 2009; for a recent review of spatial econometric OD-flow models, see Thomas-Agnan and LeSage, 2021), we propose a spatial OD threshold Tobit model to allow for a censored and directed dyadic dependent variable, of which trade data represent a typical example. We apply this new method to a cross-section of bilateral export flows to address potential multiple sources of spatial dependence in trade flows and zero trade values. Since maximum likelihood estimation (MLE) becomes infeasible in this case, we develop a Bayesian procedure to estimate the model.

Our empirical results provide supporting evidence for exporter-based (origin-based), importer-based (destination-based) and exporter-to-importer-based (origin-to-destination-based) spatial correlation in export flows. Economic sizes and geographical distance are statistically meaningful determinants of trade flows, although magnitudes of their impacts are not close to unity after taking spatial dependence into account. Besides, sizable network effects are detected for GDP, the omission of which obscures the mechanism through which economic size affects trade flows.

Model specification

LeSage and Pace's (2008) spatial OD model represents a more tailored approach to dealing with spatial correlation in data featuring an origin-to-destination flow, such as bilateral trade flows. However, their model specification is not readily applicable to the gravity model of trade if the data contain zero trade values, as their spatial OD model is proposed for continuous dependent variables. When zero trade flows are present in a data set, the presence of a limited dependent variable normally requires a limited dependent variable approach, such as a Tobit-type (spatial) model. Moreover, zero trade values pose another technical challenge when employing the log-linear form of the gravity model. To handle zero trade observations and spatial autocorrelation concurrently, this study employs a spatial OD model in conjunction with the threshold gravity model first introduced by Eaton and Tamura (1994) [1].

According to the threshold gravity model, the volume of trade between a pair of countries records a positive value only if the potential trade exceeds a certain minimum amount (i.e. threshold). As Ranjan and Tobias (2007) point out, threshold Tobit allows us to "remain true to the mixed discrete-continuous nature of trade data" by "assign[ing] meaningful probabilities to the event of no trade" and also helps to "avoid the problem of taking the log of zero" (p. 818).

Following Eaton and Tamura (1994)'s framework, the trade flow from country j to country k is modeled as:

$$\ln\left(y_{jk}^{*}+a\right)=X_{jk}\beta+\varepsilon_{jk},\ \cdots\varepsilon_{jk}\sim N\left(0\ ,\ \sigma^{2}\right) \tag{1}$$

$$\text{where } y_{jk}=\begin{cases} y_{jk}^{*}\ , & if\ \ y_{jk}^{*}>0 \\ 0\ \ , & if\ y_{jk}^{*}\leq 0 \end{cases} \quad \text{and} \quad a>max\left[0,-y_{jk}^{*}\right]$$

Or equivalently,

$$\underbrace{y_{jk}^{*}}_{\text{desired trade}}=\underbrace{-\,a}_{\text{fixed cost}}+\underbrace{\exp(X_{jk}\beta+\varepsilon_{jk})}_{\text{potential trade}} \tag{2}$$

From an economic point of view, the threshold parameter a can be interpreted as a fixed or average cost of international trade [2], and y_{jk}^* represents the *desired* amount of bilateral trade. The actual *observed* trade volume, y_{jk} , equals y_{jk}^* if the potential trade more than covers the fixed cost. Contrarily, if the potential trade falls below the fixed cost and bilateral trade becomes undesirable or unprofitable (i.e. desired trade is negative), then the *observed* trade volume y_{jk} is zero. That is to say, trade will occur only when trade is desired (when $y_{jk}^* > 0$). In this sense, the practice of arbitrarily adding "1" to all sampled trade data to force the logged term to take a defined value is not convincing, as it imposes an arbitrary one-unit trade cost.

Also, technically, the log-linear formulation of the model may help alleviate possible heteroskedasticity, as it is known that homoskedasticity in logs allows a reasonable heteroskedasticity in levels (e.g. LeSage and Thomas-Agnan, 2012).

Using matrix notation, equation (1) can be rewritten as:

$$\ln\left(y^* + a \otimes \iota_N\right) = \boldsymbol{X}\beta + \varepsilon, \cdots \varepsilon \sim N\left(0, \quad \sigma^2 \boldsymbol{I}\right) \tag{3}$$

LeSage and Pace (2008)'s spatial OD model employs three spatial lag terms of the dependent variable defined through the weight matrices $\mathbf{W}_d$, $\mathbf{W}_o$, and $\mathbf{W}_w$ to model, respectively, spatial correlation stemming from neighboring relationships among exporting countries, among importing countries, as well as concurrent neighboring relationships across trading pairs. To economize space, we do not provide the details of how to construct weight matrices $\mathbf{W}_d$, $\mathbf{W}_o$ and $\mathbf{W}_w$.(see Luo and Miller, 2014; LeSage and Pace, 2008). As explained above, the left hand side of equation (3) represents the logged value of potential trade. If spatial effects are present, we expect them to influence the latent trade volumes. Taking the log of the dependent variable attempts to correct for potential heteroskedasticity. Therefore, the spatial OD modeling of the threshold Tobit is specified as follows:

$$\begin{aligned} \ln\left(y^* + a \otimes \iota_N\right) = \rho_d \mathbf{W}_d \otimes \ln\left(y^* + a \otimes \iota_N\right) + \rho_o \mathbf{W}_o \otimes \ln\left(y^* + a \otimes \iota_N\right) \\ + \rho_w \mathbf{W}_w \otimes \ln\left(y^* + a \otimes \iota_N\right) + \boldsymbol{X}\beta + \varepsilon \end{aligned} \tag{4}$$

For brevity, we use v^* to denote $\ln(y^* + a \otimes \iota_N)$ and accordingly, (4) can be rearranged as:

$$(\mathbf{I}_\mathrm{N} - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w) \otimes v^* = \boldsymbol{X}\beta + \varepsilon \tag{5}$$

Setting $\boldsymbol{A} = (\mathbf{I}_\mathrm{N} - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)$, we can simplify equation (5):

$$\boldsymbol{A} \otimes v^* = \boldsymbol{X}\beta + \varepsilon \tag{6}$$

or equivalently,

$$v^* = \boldsymbol{A}^{-1}\boldsymbol{X}\beta + \boldsymbol{A}^{-1}\varepsilon \tag{7}$$

Eaton and Tamura (1994) rely on maximum likelihood for the estimation of their threshold Tobit model. However, once spatial lags are introduced into the model, MLE becomes infeasible, as detailed in the first part of the Supplemental Material [3]. Due to space limit, the Supplemental Material shows how we develop a Bayesian estimation algorithm which accounts for the discrete-continuous feature of trade data while avoiding the computational difficulty associated with an ML estimator.

Data, construction of weight matrices and effects estimates

(1) The data

To illustrate spatial effects among bilateral trade flows, this study employs the basic setup of the gravity equation and fits the spatial OD threshold Tobit to a sample of 32 Asian countries in 1990 (see Table A1). The explanatory variables, $\boldsymbol{X}$, include real GDP of two economies in a pair, the distance within each pair, as well as a contiguity dummy [4]. The dependent variable is bilateral exports. We collect data on export volumes, GDP and bilateral distance from Santos Silva and Tenreyro (2006)'s dataset [5]. The original data measures trade volumes in thousands of US dollars and GDP in level term. For computational manageability, this study records export volumes in billions of US dollars and GDP in millions of US dollars.

When Santos Silva and Tenreyro's dataset includes a contiguity variable, it is based only on land borders. This operationalization may be too restrictive because if two countries are separated by only a small body of water, they are *de facto* neighbors. For this reason, this study adopts a broader definition of contiguity which acknowledges not only land borders but also water borders. We retrieve the contiguity data from the Expected Utility Generation and Data Management Program (EUGene) <https://eugenesoftware.la.psu.edu/> Version 3.204, which allows the user to choose from several different distances over water under which two countries are to be considered contiguous. This study uses a conservative separation distance of less than 25 miles of water body as an alternative criterion for determining contiguity.

(2) Handling weight matrices

- *Choice of the weight matrices*

In spatial analysis, we use the weight matrices to capture inherent spatial correlation in the data. Therefore, we should construct these matrices considering the characteristics of data under study to make them more relevant to embedded spatial patterns.

Porojan (2001, p. 271) utilizes a contiguity weight matrix that codes countries that share a land border or are separated by a small body of water as contiguous. This choice is better than a predetermined capital distance as the cutoff for denoting neighbors, especially for trade participants with large territories, such as China and Russia, which might require quite large cutoff distances in order to capture possible spatial interactions. On the other hand, Porojan's weight matrix does not reflect the dyadic and directional features of flow data by differentiating the sources of spatial correlation.

As pointed out by Behrens *et al.* (2012), the weight matrix may be defined in different ways. Behrens *et al.*'s weight matrix is called the "interaction matrix" to distinguish it from the more common, distance-based formulation, because it assigns weights based on the population ratio of the exporting country in a trading pair to the total population of countries in the sample. However, this way of defining the interdependence structure of trade flows rigidly sets the influence of one trade flow invariable with respect to all other relevant trade flows. For example, if two trading pairs share the same importing country and have a similar population size in the respective exporting country (i.e. same ratio to the total population of the sample) but are differentiated by the bilateral distance within each pair, then all other bilateral trade flows involving the same importing country are supposed to exert the exactly same effect on the trade volumes of the stated two pairs according to their interaction matrix. Further, though Behrens *et al.* acknowledge that trade data features an origin-to-destination flow, their model is incapable of handling the complex connectivity structure embodied in such directional flows.

Geographic distance matters to trade behavior in different ways. Admittedly, transportation cost is a factor when countries decide with whom to trade. Yet, physical proximity allows countries to benefit from the spread of technologies, ideas and policies, which are all conducive to the promotion of trade (e.g. technology spillover, LeSage *et al.*, 2007). Therefore, we still rely on a first-order contiguity weight matrix $\mathbf{W}$, where $\mathbf{W}_{ij}$ $(i \neq j)$ is coded as "1" if the two members of a pair are contiguous and "0" otherwise. We include a row-standardized weight matrix $\mathbf{W}$ based on the stated definition of contiguity in Table A2. We then adapt this $\mathbf{W}$ matrix to the neighboring relationships unique to OD flows to build the three spatial weight matrices $\mathbf{W}_d$, $\mathbf{W}_o$ and $\mathbf{W}_w$ to specifically model origin-centric, destination-centric and origin-to-destination-centric dependence as proposed by LeSage and Pace (2008). Through Kronecker product operations, $\mathbf{W}_d = \mathbf{I}_n \otimes \mathbf{W}$, reflect that factors causing trade flows from an exporting country to an importing country may induce or dampen similar flows to nearby importers; $\mathbf{W}_o = \mathbf{W} \otimes \mathbf{I}_n$ reflects that origin-based dependence wherein a country's exporting behavior may simulate or impede similar flows of trade from its neighboring exporters to the same destination; and $\mathbf{W}_w = \mathbf{W} \otimes \mathbf{W}$ represents a second-order connectivity between the neighborhood of an exporting country and the neighborhood of an importing country [6]

- *Elimination of self-directed pairs*

In keeping with the structure of weight matrix $\mathbf{W}$, by construction the data include self-directed pairs. However, when the values of the dependent variable are set to zero for self-directed pairs, this operation fails to distinguish the zero values for these observations from those observed for no-trade pairs, thus leading to biased estimation. To avoid this bias, several researchers include an additional set of explanatory variables to estimate a model of self-directed pairs in tandem with a model of bilateral flows (see LeSage and Pace, 2008).

Behrens *et al.* (2012) include internal absorptions in their data; however, their construction of the "interaction matrix" assigns no weights to self-directed pairs (i.e. own trade flows), thus excluding the role of self-directed pairs from the interactive network. Santos Silva and Tenreyro (2006) consider only non-self-directed trade pairs in their investigation of the proper functional form for the gravity equation. Accordingly, it is difficult to assume that bilateral trade flows and internal flows operate under the same mechanism and should be estimated jointly, especially for the cross-section considered in this study's application. In 1990, external trade volumes and internal trade flows were not on a comparable scale for many of the sample countries, either because they were not yet capable of participating in international trade or because they focused on an inward looking trade policy with international trade accounting for only a small part of national income. In this circumstance, it may not be necessary to include self-directed pairs in the model. However, to accommodate diverging perspectives, this study fits the spatial OD threshold Tobit model to data augmented with internal flows as an empirical exercise [7]. The latter case does not require the elimination process as described below, which removes the structural rigidity of the weight matrices associated with self-directed dyads.

To eliminate the role of self-directed pairs from the estimation, the $(n(m-1)+m)$ th rows and columns of the weight matrices need to be removed, with n denoting the number of countries in the sample and $m = 1,\ 2,\ \cdots,\ n$. This can be accomplished by pre-multiplying each weight matrix by a "selection" matrix $\boldsymbol{B}$ and post-multiplying the resulting matrix by the transpose of $\boldsymbol{B}$. Here, $\boldsymbol{B}$ represents an $n(n-1)$ by n^2 sparse matrix. The elimination procedure removes self-directed pairs, freeing the researcher from inadvertent constraints innate to the construction of spatial weight matrices. Using matrix notation, this process is expressed as

$$\boldsymbol{B}v^* = \rho_d \boldsymbol{B}\mathbf{W}_d^R \boldsymbol{B}'\boldsymbol{B}v^* + \rho_o \boldsymbol{B}\mathbf{W}_o^R \boldsymbol{B}'\boldsymbol{B}v^* + \rho_w \boldsymbol{B}\mathbf{W}_w^R \boldsymbol{B}'\boldsymbol{B}v^* + \boldsymbol{B}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{B}\varepsilon, \tag{20}$$

where v^* is defined as above. $\mathbf{W}_d^R$, $\mathbf{W}_o^R$ and $\mathbf{W}_w^R$ signify a renormalization of the modified weight matrices excluding neighboring relationships with self-directed pairs, though renormalization takes place after the elimination step.

- *Weight matrices compliant with the restructured data*

As explained earlier, we need to rearrange the data to stack zero-valued observations above non-zero ones in order to take advantage of the properties of a multivariate normal distribution and derive the conditional density of latent v_1^* given v_2. These data restructuring necessitate additional operations on the abovementioned modified weight matrices $\mathbf{W}_d^R$, $\mathbf{W}_o^R$ and $\mathbf{W}_w^R$ so as to keep the inherent connectivity structure intact. However, this procedure is more data-specific, depending on the positions of zero observations in the data.

Let $[\boldsymbol{D}]_i$ represent the rows of an N by N identity matrix $\boldsymbol{D}$ with the row numbers, i, corresponding to the positions of zero observations in the data that already have self-directed pairs removed. Then place the block of $[\boldsymbol{D}]_i$ above the remaining rows of $\boldsymbol{D}$, denoted as $[\boldsymbol{D}]_{i-}$, to create a new matrix $\boldsymbol{M}$. That is $\boldsymbol{M} = ([\boldsymbol{D}]_i', [\boldsymbol{D}]_{i-}')'$. Pre-multiplying each of the matrices, $\mathbf{W}_d^R$, $\mathbf{W}_o^R$ and $\mathbf{W}_w^R$, by $\boldsymbol{M}$ and then post-multiplying each by the transpose of $\boldsymbol{M}$ produce revised weight matrices complying with the restructured data (i.e. zero values placed on top of non-zero observations). Hence, the new model is given by

$$\begin{aligned}\boldsymbol{M}(\boldsymbol{B}v^*) &= \rho_d \boldsymbol{M}(\boldsymbol{B}\mathbf{W}_d^R\boldsymbol{B}')\boldsymbol{M}'[\boldsymbol{M}(\boldsymbol{B}v^*)] + \rho_o \boldsymbol{M}(\boldsymbol{B}\mathbf{W}_o^R\boldsymbol{B}')\boldsymbol{M}'[\boldsymbol{M}(\boldsymbol{B}v^*)] \\ &\quad + \rho_w \boldsymbol{M}(\boldsymbol{B}\mathbf{W}_w^R\boldsymbol{B}')\boldsymbol{M}'[\boldsymbol{M}(\boldsymbol{B}v^*)] + \boldsymbol{M}(\boldsymbol{B}\boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{M}(\boldsymbol{B}\varepsilon)\end{aligned} \tag{21}$$

Given that $\boldsymbol{M}$ is an orthogonal matrix, (21) can be simplified as [8]

$$\begin{aligned}\boldsymbol{M}(\boldsymbol{B}v^*) &= \rho_d \boldsymbol{M}(\boldsymbol{B}\mathbf{W}_d^R\boldsymbol{B}')(\boldsymbol{B}v^*) + \rho_o \boldsymbol{M}(\boldsymbol{B}\mathbf{W}_o^R\boldsymbol{B}')(\boldsymbol{B}v^*) \\ &\quad + \rho_w \boldsymbol{M}(\boldsymbol{B}\mathbf{W}_w^R\boldsymbol{B}')(\boldsymbol{B}v^*) + \boldsymbol{M}(\boldsymbol{B}\boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{M}(\boldsymbol{B}\varepsilon)\end{aligned} \tag{22}$$

(3) Marginal effects in spatial OD Tobit models

It is not straightforward to interpret the estimated coefficients of a Tobit model due to its inherent nonlinearity. The spatial autoregressive structure of the spatial OD threshold Tobit model further complicates interpretation. Using marginal effects, partial derivatives reflecting how changes in an explanatory variable affect the expected value of y_i is useful. As in other spatial OD models, the origin-centric ordering of the variables in a spatial OD Tobit implies that, for non-bilateral regressors (e.g. GDP in this application), marginal effects should be calculated as country-specific rather than dyad (or observation)-specific, because a change in one country's regressor immediately affects all dyads in which that country is either an origin or a destination and then the effects are propagated through the spatial spillover mechanism to other dyads.

Given that interpretation of coefficient estimates of a conventional regression model averages over impacts on all observations arising from changes in explanatory variables, LeSage and Thomas-Agnan (2012) propose using scalar summary measures providing interpretation of spatial autoregressive interaction models consistently. By averaging over the relevant marginal effects associated with changing a given characteristic for all regions, $i = 1, \ldots, n$, these scalar summaries allow the calculation of direct effects – i.e. origin and destination effects arising from changing a typical country's regressor on pairs involving that country, distinguished by the origin or destination status of the said country in those pairs, network (indirect) effects – i.e. the effects on pairs not involving the said country, as well as intraregional effects on self-directed pairs. Total effects are the sum of these four types of effects.

In matrix notation, the partial derivatives measuring total effects on the latent variable (represented by the flow matrix Y^*) from changing X_i^k (region $i = 1, \ldots, n$ and characteristic $k = 1, \ldots, K$) are given by

$$\mathrm{TE} = \begin{pmatrix} \partial Y^*/\partial X_1^k \\ \partial Y^*/\partial X_2^k \\ \vdots \\ \partial Y^*/\partial X_n^k \end{pmatrix} = (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1} \begin{pmatrix} Qd_1\beta_d^k + Qo_1\beta_o^k \\ Qd_2\beta_d^k + Qo_2\beta_o^k \\ \vdots \\ Qd_N\beta_d^k + Qo_N\beta_o^k \end{pmatrix} \quad (23)$$

where Qd_i is an $n \times n$ matrix of zeros with the ith row adjusted to be a vector of ones, and Qo_i is an $n \times n$ zero matrix with the ith column replaced by a vector of ones. A scalar summary measuring the total effects of a change in the typical region's kth characteristic can be obtained by averaging across all the elements of the $N \times n$ matrix TE in (23) and thus takes the form: $te = (1/N)l'_N \otimes TE \otimes l_n$.

We calculate a scalar summary of the destination effects by averaging across the elements in the matrix TE corresponding to the partial derivatives for pairs in which the country with changed characteristic is the destination. Mathematically, we express this scalar measure as $de = (1/N)l'_N \otimes DE \otimes l_n$, where DE is an $N \times n$ matrix that retains the $[1 + n \otimes (r-1)]$ th rows ($r = 1, \ \ldots, \ n$) of TE matrix while setting the remaining rows to zero.

Similarly, we construct a scalar summary of the origin effects by averaging across the elements in the matrix TE corresponding to the partial derivatives for pairs in which the country with changed characteristic is the origin, expressed as $oe = (1/N)l'_N \otimes OE \otimes l_n$, where OE is an $N \times n$ matrix that retains the $[(1 + n \otimes (r-1)) : nr, r]$ elements of TE (i.e. the rth n elements of the rth column of the matrix TE with $r = 1, \ \ldots, \ n$) while setting the remaining elements to zero.

Further, we create a scalar summary of the intraregional effects using $ie = (1/N)l'_N \otimes IE \otimes l_n$, with IE as an $N \times n$ zero matrix adjusted to pass over, in the corresponding row and column positions, the elements of the matrix TE that represent partial derivatives for self-directed pairs.

Consequently, we obtain a scalar summary for the network effects using: $ne = te - de - oe - ie$.

In order to interpret the newly proposed spatial OD threshold Tobit model, this study adopts LeSage and Thomas-Agnan's approach with a modification that sets intraregional effects to 'zero' when internal flows are not included in the model estimation. Not surprisingly, the nonlinear nature of Tobit means that the β_d^k and β_o^k in (23) should be replaced by derivatives that are no longer constant scalars as would be the case of a linear regression model. Instead, they are two $n \times n$ matrices with varying elements depending on the specific values of explanatory variables observed for each OD flow. This study uses bold letters to distinguish them from coefficient estimates. By organizing the column vector $\boldsymbol{X\beta}$ into an $n \times n$ origin-centric matrix $\boldsymbol{Z}$, we write [9].

$$\boldsymbol{\beta_d^k} = \Phi\left(\frac{\boldsymbol{Z} - \boldsymbol{H} \otimes \tau}{\boldsymbol{U}}\right) \otimes \beta_d^k + \ \phi\left(\frac{\boldsymbol{H} \otimes \tau - \boldsymbol{Z}}{\boldsymbol{U}}\right) \otimes \left(-\frac{\beta_d^k}{\boldsymbol{U}}\right) \otimes \tau$$

and

$$\boldsymbol{\beta_o^k} = \Phi\left(\frac{\boldsymbol{Z} - \boldsymbol{H} \otimes \tau}{\boldsymbol{U}}\right) \otimes \beta_o^k + \ \phi\left(\frac{\boldsymbol{H} \otimes \tau - \boldsymbol{Z}}{\boldsymbol{U}}\right) \otimes \left(-\frac{\beta_o^k}{\boldsymbol{U}}\right) \otimes \tau \quad (24)$$

where $\boldsymbol{H}$ stands for an $n \times n$ matrix of ones, $\boldsymbol{\tau}$ represents the censoring point, and $\boldsymbol{U}$ designates an $n \times n$ diagonal matrix with the diagonals set equal to the square root of the diagonals in the covariance matrix $\boldsymbol{\Omega}$ defined in (16).

Empirical results and interpretation

Following the algorithm described in section (b), this study runs 35,000 iterations. Inspection of the trace plots for all model parameters indicates quick convergence to a steady state. Thus, this research uses a burn-in period of 5,000 iterations and draws inferences based on the remaining 30,000 iterations. As is conventional practice in Bayesian analysis, a 95% credibility interval together with posterior mean and standard deviation associated with each model parameter are reported in Table 1.

Table 2 displays the results of several other techniques commonly used for the estimation of the gravity equation alongside those from the spatial OD threshold Tobit [10]. Column 1 presents ordinary least squares (OLS) estimates using the logarithm of exports as the dependent variable. As noted earlier, this requires dropping all observations of zero bilateral trade flow. Only 612 country pairs, or 61.7% of the current sample, record positive export flows. Column 2 shows the OLS estimates with $\ln(y_i + 1)$ as the dependent variable, and Column 3 presents OLS results using $\ln(y_i + 0.0049)$ as the dependent variable, where the added positive constant is opted in light of the threshold estimate from the threshold Tobit model (see Column 5). Column 4 exhibits results of standard Tobit and Column 5 lists threshold Tobit estimates based on Eaton and Tamura (1994). Column 6 presents the spatial OD threshold Tobit results in a compatible format.

The signs of all of the parameter estimates are remarkably stable across all models except for the contiguity variable, which seems not substantially different from zero in these models. Inspecting the first three columns, we find that different approaches to log transforming the dependent variable lead to noticeable changes in OLS estimates. As shown in Column 1, the coefficients for exporter's GDP and distance are almost equal to positive one and negative one, respectively, while the GDP coefficient for importer is also on a comparable scale. However, we obtained these results using positive export flows only. When we include the zero observations for estimation, the magnitude of conventional trade variables decreases noticeably. As illustrated in Column 3, when we set the added positive constant (i.e. fix the threshold parameter) as 0.0049, the sizes of the two income elasticities decrease by more than half, with exporter income-elasticity declining from 0.9970 to 0.4404 and importer income-elasticity decreasing from 0.8813 to 0.4013. As for the distance parameter, its magnitude changes from −1.0081 to −0.2006. Further, when we add an arbitrary constant of "1" to the export flow data before log transforming them as in some previous trade studies, the sizes of all parameters except for contiguity decrease to about one-twentieth of those estimated only with positive observations [11]. Once a threshold parameter is included, the sign of the coefficient on contiguity changes from negative to positive as shown in Columns 5 and 6, compared to the others, which is consistent with predictions of trade theory on bilateral trade costs. Though this coefficient seems not significantly different from zero in almost all the models. Also, the coefficient is quite consistent across these models. The magnitude of coefficient on exporter's income is larger than that of importer's; though this coefficient estimate itself is not directly comparable across the models, which we will discuss below.

	Mean	S.D.	2.5%	Median	97.5%	Sample
Intcpt	−6.7752	0.6666	−8.0859	−6.7740	−5.4725	30,000
Log exporter's GDP	0.3286	0.0270	0.2767	0.3285	0.3816	30,000
Log importer's GDP	0.2877	0.0268	0.2359	0.2876	0.3410	30,000
Log distance	−0.1842	0.0660	−0.3125	−0.1839	−0.0558	30,000
Contiguity	0.0331	0.1876	−0.3324	0.0333	0.4005	30,000
ρ_d	0.3498	0.0299	0.2905	0.3499	0.4078	30,000
ρ_o	0.3418	0.0323	0.2793	0.3420	0.4069	30,000
ρ_w	−0.1473	0.0371	−0.2221	−0.1476	−0.0731	30,000
α	0.0049	0.0001	0.0047	0.0049	0.0050	30,000

Table 1. Bayesian estimates of the spatial OD threshold Tobit model of export Flows, Asia, 1990

Table 2.
Regression estimates of the traditional gravity equation

Estimator Dep. Var.	OLS $\ln(y_i)$	OLS $\ln(y_i+1)$	OLS $\ln(y_i+0.0049)$	Tobit $\ln(y_i)$	Threshold Tobit $\ln(y_i+a)$	Spatial OD threshold Tobit[a] $\ln(y_i+a)$	Spatial OD threshold Tobit[b] $\ln(y_i+a)$
Intcpt	−15.8456 *	−0.6689 *	−10.7401*	−15.8567 *	−15.6865 *	−6.7752 *	−8.8987 *
	(−18.2169, −13.4744)	(−0.9225, −0.4153)	(−11.8838, −9.5964)	(−18.2218, −13.4916)	(−17.4427, −13.9303)	(−8.10859, −5.4725)	(−10.4636, −7.4026)
Log exp-GDP	0.9970 *	0.0559 *	0.4404 *	0.9979 *	0.6473 *	0.3286 *	0.4920 *
	(0.9084, 1.0856)	(0.0477, 0.0641)	(0.4030,0.4778)	(0.9095, 1.0863)	(0.5885, 0.7061)	(0.2767, 0.3816)	(0.4278, 0.5586)
Log imp-GDP	0.8813 *	0.0594 *	0.4013 *	0.8812 *	0.5740 *	0.2877 *	0.4638 *
	(0.7956, 0.9670)	(0.0512, 0.0676)	(0.3639, 0.4387)	(0.7957, 0.9667)	(0.5187, 0.6293)	(0.2359, 0.3410)	(0.4017, 0.5310)
Log distance	−1.0081 *	−0.0442 *	−0.2006 *	−1.0079 *	−0.1907 *	−0.1842 *	−0.9395 *
	(−1.2925, −0.7237)	(−0.0740, −0.0144)	(−0.3351, −0.0661)	(−1.2917, −0.7241)	(−0.3536, −0.0278)	(−0.3125, −0.0558)	(−1.0535, −0.8297)
Contiguity	−0.0982	−0.0931 *	−0.0479	−0.0967	0.1517	0.0331	−0.4162
	(−0.8308, 0.6344)	(−0.1786, −0.0076)	(−0.4334, 0.3376)	(−0.8274, 0.6340)	(−0.3069, 0.6103)	(−0.3324, 0.4005)	(−0.8763, 0.0427)
ρ_d						0.3498 *	0.1514 *
						(0.2905, 0.4078)	(0.0917, 0.2132)
ρ_o						0.3418 *	0.1131 *
						(0.2793, 0.4069)	(0.0472, 0.1748)
ρ_w						−0.1473 *	0.0512
						(−0.2221, −0.0731)	(−0.0209, 0.1153)
α					0.0049 *	0.0049 *	0.0049 *
					(0.0033, 0.0065)	(0.0047, 0.0050)	(0.0048, 0.0050)

Note(s): 95% confidence intervals for OLS and ML-based estimates, and 95% credible intervals for Bayesian estimates (sample size of 30,000)
* denotes zero not in interval
[a] considers bilateral trade flows only ($N = 992$), i.e. excluding self-directed pairs
[b] includes both bilateral and internal trade flows ($N = 1024$)

Estimates from spatial OD threshold Tobit suggest that bilateral trade flows are correlated in space and the interdependence among export flows arises from multiple sources. Specifically, none of the 95% credible intervals for spatial coefficients contains the value zero, with ρ_d and ρ_o both showing a positive sign and ρ_w negative. The positive sign of ρ_d suggests that when a country exports to another country, it is likely to export to the neighbors of its destination as well. This spillover effect may be partly due to potential economies of scale. Exporting to countries that are clustered geographically allows an exporter to take advantage of the established trade route and existing infrastructures of export activities. On the demand side, countries that are in close geographic proximity, especially smaller ones, are predisposed to a similar endowment of resources, which may lead them to import the same types of goods. A positive ρ_o signals that one country's exports to a given destination tend to be positively related to trade volumes from neighboring countries to the same importer, a different type of spillover effect. This resemblance in trading behavior among geographically proximate exporters may be attributable to the ease of dissemination of technologies and innovations, relocation of skilled labor and even policy diffusion occurring among neighboring countries. In this sense, proximity appears to provide trade-promoting opportunities rather than create market competition. Moreover, similarities in resource endowment among neighboring exporting countries may lead to specializing in the production of same or similar goods.

On the other hand, the negative sign of ρ_w indicates a competitive relationship across trading pairs when a 'dual' neighboring relationship exists at both origins and destinations. When both exporter countries eye the same export markets while both importer countries look to the same suppliers, stronger trade ties within one pair may cause the other exporter (or importer) to concern about disadvantaged trade position with the importer (or exporter) in the said pair. In the context of concurrent neighboring relationships across two pairs, between both exporters and importers, this competition is likely to induce a negative impact on trade flows within the other pair.

The estimate of the threshold parameter is around 0.0049, and zero falls outside the 95% credible intervals. This implies that, on average, potential trade volumes need to be at least 4.9 million for an exporter country to be willing (i.e. for it to be profitable) to trade.

It appears that geographic distance negatively affects trade volume. Its coefficient estimate from the spatial OD threshold Tobit is −0.1842. However, this estimate is different from those obtained under standard OLS and Tobit models (columns 1 and 4), which are very close to unity, −1.0081 and −1.0079, respectively. It is also slightly smaller than the distance coefficient produced by Eaton and Tamura's threshold Tobit. This is consistent with LeSage and Thomas-Agnan (2012)'s observation that the importance of distance diminishes after accounting for spatial dependence "often . . . for the spatial variants of gravity models" (p. 23). In a similar vein, Fotheringham and Webber (1980) note that in the presence of spatial autocorrelation, the estimated parameter on the distance variable captures both "a 'true' friction of distance effect" and a measure of the map pattern (p. 34). Joining their insight, Porojan (2001, p. 275) further explicates that spatial lag in his model captures an important part of the spatial effect, which the traditional formulation of the gravity model partially picked up through the distance variable. Since spatial OD modeling is better tailored to flow data in capturing spatial effects, it is expected that the estimated impact of distance weakens.

Contiguity shows no discernible impact on export flows. Although the respective coefficient takes a positive sign for both the spatial and non-spatial threshold Tobit, zero falls near the center of the intervals for this coefficient in all but one model. This is consistent with Ranjan and Tobias (2007)'s finding. While the authors do not offer a formal explanation for the insignificance of the contiguity effect, they draw attention to the difference in their new model specification which accounts for the discrete-continuous nature of bilateral trade data (p. 830). More importantly, in competition with spatial terms built on contiguity relationship, a bilateral contiguity dummy may prove inadequate to distinguish the involved effects of contiguity on trade flows.

Whether the spatial OD threshold Tobit model is estimated with or without observations on internal trade flows (i.e. columns 6 and 7), the estimation results are consistent in the sign and significance of coefficients. According to the design of spatial weight matrix **W**, we set all diagonal elements (i.e. weights assigned to intra-regional flows or in this application, internal flows) to zero. This structure is transferred to $\mathbf{W}_d$, $\mathbf{W}_o$ and $\mathbf{W}_w$, leading to different treatment of neighboring relationships with self-directed and non-self-directed pairs. Thus, the inclusion of own trade flows may attenuate spatial effects. Nonetheless, positive exporter- and importer-based dependence still emerge for this extended data, though the magnitude of the coefficient estimates is smaller. Besides, the coefficient size on *distance* increases as expected. Compared to bilateral trade flows, internal flows record much larger trade volumes, but occur within relatively shorter distances, tilting toward a resisting effect of distance on trade. Moreover, both GDP coefficients are estimated as larger with the addition of internal flow data, reflecting that internal flows account for a much higher proportion of GDP.

As LeSage and Thomas-Agnan (2012) point out, estimates for non-bilateral variables in spatial OD models (i.e. exporter-GDP and importer-GDP in the current application) are not directly comparable to those from OLS. Hence, it is more appropriate to calculate scalar summary effect estimates that reflect marginal effects associated with changes in regional characteristics on average flows that provide interpretation consistent with conventional linear regression models. Table 3 shows summary effect estimates of GDP for the spatial OD threshold Tobit model. The first column displays (averaged) marginal effects on the latent y_i^*, while the second column presents (averaged) marginal effects on y_i. The scalar summary estimates for the latent and observed regressands are similar. For the effect estimates for y_i, a one percent increase in GDP of the typical exporter (i.e. origin) country likely leads to a 0.5941% increase in export flows, while a one percent increase in GDP of the typical importer (i.e. destination) country likely leads to a 0.3351% increase in export flows. These results imply that the impact of exporter's GDP is larger than that of importer's GDP. The model also estimates network effects to be positive, suggesting that a one percent increase in the GDP of the typical country likely leads to a 0.2190% increase in export flows due to spatial spillover (separately from flows already captured by the origin and destination effects). In this example, total effects reflect the sum of origin, destination and network effects on trade flows. The total effect of increasing GDP by one percent is a 1.1482% increase in export flows.

As an exploratory step, Table 4 compares the marginal effects of GDP for the latent variable y^* in the spatial Tobit as well as the three OLS models considered in this study. This comparison addresses whether the different coefficients in these models are compensated for

Table 3.
Marginal estimates of GDP for the spatial OD threshold Tobit

	Marginal effects on y*	Marginal effects on y
Origin effects	0.5180	0.5941
Destination effects	0.2877	0.3351
Network effects	0.1993	0.2190
Total effects	1.0050	1.1482

Table 4.
Effect estimates of GDP for the OLS models and for the latent variable in the spatial OD threshold Tobit

Estimator Dep. Var	OLS $\ln(y_i)$	OLS $\ln(y_i + 1)$	OLS $\ln(y_i + 0.0049)$	Spatial OD threshold Tobit $\ln(y_i^* + a)$
Origin effects	0.9970	0.0559	0.4404	0.5180
destination effects	0.8813	0.0594	0.4013	0.2877
Network effects	0	0	0	0.1993
Total effects	1.8783	0.1153	0.8417	1.0050

by the different a's, producing similar marginal effects. For illustration, we focus on the application that considers bilateral flows only (i.e. columns 1–6).

For models with no spatial correlation (i.e. the first three columns), the origin effects and destination effects are the same as the coefficient estimates for exporter-GDP and importer-GDP and the network effects are zeros. As shown in Table 4, the marginal effects from these different models are dissimilar, though all four types of summary measures consistently identify positive impacts of GDP on export flows as expected. This implies that the different impacts estimated of GDP by the spatial OD threshold model are not statistical artifacts of the choice of a. By allowing for spatial correlation, spatial OD threshold Tobit detects sizeable network effects, estimated as 0.1993, when examining the effects on the latent variable. The asymmetric income impacts between exporter and importer are more distinct under the spatial model, with origin effects (0.5180) about 1.8 times the magnitude of destination effects (0.2877). Nonetheless, both origin and destination effects are smaller than unity, revealing a much reduced influence of GDPs on export volumes once spatial dependence is appropriately controlled. Several previous studies discuss this issue. For instance, Grossman (1998) questions the unrealistically large magnitude of coefficients on GDPs. Porojan (2001) reports considerable changes in the size of estimated parameters when applying alternative spatial econometric models to both import and export data, though he does not calculate marginal effects, which should provide more appropriate interpretation given the nonlinearity introduced by spatial correlation. Moreover, applied trade economists have always been aware of potential omitted variables in the empirical specification of the gravity model. If the model leaves out variables which correlate with GDP measures and positively affect export activities, it will inflate estimates of GDP coefficients. Several studies recommend spatial lags as a more efficient way of dealing with omitted variables (LeSage and Pace, 2008; Behrens *et al.*, 2012; Porojan, 2001). For instance, Behrens *et al.* (2012) argue that using lagged terms is "more robust to potential misspecification concerning the form of interdependence" (p. 775).

As the econometric models in Table 2 have different assumptions of underlying data distribution and are estimated using different techniques, model comparison is not straightforward. Thus, the second part of the Supplementary Material employs several measures of model fit as exploratory tools, which are helpful in evaluating how well the model represents the data [3].

Conclusion

We make two contributions in this study. Methodologically, we advance an econometric model by considering the complexity of spatial autocorrelation embedded in "directional" trade flows, while dealing with the corner solution where trade volumes are zero. Empirically, we provide evidence that bilateral trade flows are correlated in space and that conventional trade variables have lesser impact than previously reported, working through multiple channels due to multifaceted spatial dependence of trade flows. On the other hand, in fitting the sample data that contain both a sizable number of zeros as well as some particularly large values, we notice that our spatial OD threshold Tobit model performs better than the non-spatial threshold Tobit model. Future research should try to further improve the model by addressing issues related to zero and extreme trade values.

Notes

1. Applying a Bayesian procedure to the threshold gravity model proposed by Eaton and Tamura (1994), Ranjan and Tobias (2007) examine the impact of contract enforcement on bilateral trade flows. LeSage and Pace (2009) briefly mention the potential to combine their spatial Tobit model with the threshold value of trade idea proposed by Eaton and Tamura and later adopted by Ranjan and Tobias.
2. Similarly, Rauch and Trindade (2002, p. 119) think of a as "an amount of 'melting' that occurs as soon as the trip starts, independent of the distance travelled." See also Ranjan and Tobias (2007).

3. The Supplemental Material is available online at https://whanchoi.people.uic.edu/research.html
4. The explanatory variables are all log transformed except the contiguity variable.
5. Santos Silva and Tenreyro post their data and definition of variables at http://privatewww.essex.ac.uk/~jmcss/LGW.html. When comparing different estimators using the Anderson and van Wincoop (2003) gravity model, which controls for multilateral resistance by including exporter- and importer-specific effects, Santos Silva and Tenreyro (2006) do not use countries' GDPs as explanatory variables as the cross-sectional data employed can only identify bilateral variables.
6. In this study, lower case n denotes the number of countries in the sample, whereas the upper case N stands for the total number of observations, which equals n^2 by construction according to the spatial OD modeling set forth by LeSage and Pace (2008), and equals $n^2 - n = n(n-1)$ after self-directed pairs are removed from the data using the elimination step proposed in the next section.
7. As a crude measure, we calculates internal trade flows as the difference between GDP and trade balance as suggested by Lebreton and Roi (2009, *p*. 5). To maintain data consistency, we multiply GDP data by external balance on goods and services (% of GDP) to back calculate trade balance. We draw data on external balance of goods and services (% of GDP) from World Development Indicators (WDI) online version. For Cambodia and the United Arab Emirates, this data is not available for 1990, we use the year this information first becomes available (1993 for Cambodia, 2001 for UAE). We take data on internal distance from the *GeoDist* database compiled by Mayer and Zignago (2011), available at cepii.fr/anglaisgraph/bdd/distances.htm. We compute the internal distance of a country as $d_{ii} = 0.67\sqrt{area/\pi}$, with country area measured in square kilometers.
8. Since $\boldsymbol{M}$ is a square matrix with orthonormal column (and row) vectors, we know $\boldsymbol{M}$ is orthogonal and $\boldsymbol{M}'\boldsymbol{M} = \boldsymbol{M}^{-1}\boldsymbol{M} = \boldsymbol{I}$.
9. To comply with the origin-centric ordering, the diagonal elements of the matrix $\boldsymbol{Z}$ in this application are set to zero as place holders, since we do not consider sample countries' internal trade flows. And subsequently, we replace the diagonals of $\boldsymbol{\beta}_{\boldsymbol{d}}^{\boldsymbol{k}}$ and $\boldsymbol{\beta}_{\boldsymbol{o}}^{\boldsymbol{k}}$ with zeros to reflect the exclusion of interregional flows.
10. For non-Bayesian estimations, a 95% confidence interval is reported in parentheses below each point estimate, while for the Bayesian estimation a 95% credible interval is presented.
11. Although in this case, the value "1" is quite large given that export flows are measured in billions of US dollars, this exercise illustrates that the choice of the positive constant added to trade data in order to make use of the log-linearized gravity equation does affect the estimation results and therefore should not be made ad hoc. For instance, Behrens *et al.* (2012) augment zero trade flows by adding 1, which might have exerted an unduly impact on their estimates given that the Canada-US exports dataset is measured in million US dollars for the year 1993.

References

Anderson, J.E. (2011), "The gravity model", *American Review of Economics*, Vol. 3, pp. 133-160.

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

Behrens, K., Ertur, C. and Koch, W. (2012), "'Dual' gravity: using spatial econometrics to control for multilateral resistance", *Journal of Applied Econometrics*, Vol. 27 No. 5, pp. 773-794.

Eaton, J. and Tamura, A. (1994), "Bilateralism and regionalism in Japanese and U.S. Trade and direct foreign investment patterns", *Journal of the Japanese and International Economies*, Vol. 8, pp. 478-510.

Feenstra, R.C. (2004), *Advanced International Trade*, Princeton University Press, Princeton.

Fotheringham, A.S. and Webber, M.J. (1980), "Spatial structure and the parameters of spatial interaction models", *Geographical Analysis*, Vol. 12 No. 1, pp. 33-46.

Grossman, G.M. (1998), "Comment on determinants of bilateral trade: does gravity work in a neoclassical world? By Alan V. Deardorff", in Frankel, J.A. (Ed.), *The Regionalization of the World Economy*, Chicago University Press, London, pp. 29-31.

Helpman, E., Melitz, M. and Rubinstein, Y. (2008), "Estimating trade flows: Trading partners and trading volumes", *Quarterly Journal of Economics*, Vol. 123 No. 2, pp. 441-487.

LeSage, J.P. and Pace, R.K. (2008), "Spatial econometric modeling of origin-destination flows", *Journal of Regional Science*, Vol. 48 No. 5, pp. 941-967.

LeSage, J.P. and Pace, R.K. (2009), *Introduction to Spatial Econometrics*, Chapman & Hall/CRC, Boca Raton.

LeSage, J.P. and Thomas-Agnan, C. (2012), *Interpreting Spatial Econometric Origin-Destination Flow Models*, Texas State University, San Marcos, TX.

LeSage, J.P., Fischer, M.M. and Scherngell, T. (2007), "Knowledge spillovers across Europe: evidence from a poisson spatial interaction model with spatial effects", *Papers in Regional Science*, Vol. 86 No. 3, pp. 393-421.

Luo, S. and Miller, J.I. (2014), "On the spatial correlation of international conflict initiation and other binary and dyadic dependent variables", *Regional Science and Urban Economics*, Vol. 44, pp. 107-118.

Mayer, T. and Zignago, S. (2011), "Notes on CEPII's distance measures", CEPII Working Paper 201-25.

McCallum, J. (1995), "National borders matter", *American Economic Review*, Vol. 85 No. 3, pp. 615-623.

Pfaffermayr, M. (2019), "Gravity models, PPML estimation and the bias of the robust standard Errors", *Applied Economic Letter*, Vol. 26 No. 18, pp. 1-5.

Porojan, A. (2001), "Trade flows and spatial effects: the gravity model revisited", *Open Economies Review*, Vol. 12, pp. 265-280.

Ranjan, P. and Tobias, J.L. (2007), "Bayesian inference for the gravity model", *Journal of Applied Econometrics*, Vol. 22, pp. 817-838.

Rauch, J.E. and Trindade, V. (2002), "Ethnic Chinese networks in international trade", *Review of Economics and Statistics*, Vol. 84 No. 1, pp. 116-130.

Santos Silva, J.M.C. and Tenreyro, S. (2006), "The log of gravity", *Review of Economics and Statistics*, Vol. 88 No. 4, pp. 641-658.

Thomas-Agnan, C. and LeSage, J.P. (2021), "Spatial econometric OD-flow models", in Fischer, M.M. and Nijkamp, P. (Eds), *Handbook of Regional Science*, Springer-Verlag, GmbH, pp. 2179-2199.

Tinbergen, J. (1962), *Shaping the World Economy*, Twentieth Century Fund, New York.

Weidner, M. and Zylkin, T. (2021), "Bias and consistency in three-way gravity models", *Journal of International Economics*, Vol. 132 No. 103513, pp. 1-22.

Appendix

Bahrain (BAH)	Jordan (JOR)	Russian Federation (RUS)
Bangladesh (BNG)	Korea, Rep. (ROK)	Saudi Arabia (SAU)
Bhutan (BHU)	Lao PDR (LAO)	Singapore (SIN)
Brunei (BRU)	Lebanon (LEB)	Sri Lanka (SRI)
Cambodia (CAM)	Malaysia (MAL)	Syrian Arab Rep. (SYR)
China (CHN)	Maldives (MAD)	Thailand (THI)
India (IND)	Mongolia (MON)	Turkey (TUR)
Indonesia (INS)	Nepal (NEP)	United Arab Emirates (UAE)
Iran (IRN)	Oman (OMA)	Vietnam (DRV)
Israel (ISR)	Pakistan (PAK)	Yemen (YEM)
Japan (JPN)	Philippines (PHI)	

Table A1.
List of sample countries (abbreviations in parentheses)

	RUS	IRN	TUR	SYR	LEB	JOR	ISR	SAU	YEM	BAH	UAE	OMA	CHN	MON	ROK	JPN	IND	BHU	PAK	BNG	SRI	MAD	NEP	THI	CAM	LAO	DRV	MAL	SIN	BRU	PHI	INS	SUM
RUS	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
IRN	1	0	1	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	7
TUR	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
SYR	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
LEB	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
JOR	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
ISR	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
SAU	0	1	0	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
YEM	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
BAH	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
UAE	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
OMA	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
CHN	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	1	0	0	1	1	0	0	0	0	0	9
MON	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
ROK	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
JPN	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
IND	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	8
BHU	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
PAK	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
BNG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
SRI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MAD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NEP	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
THI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	4
CAM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	3
LAO	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	4
DRV	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	3
MAL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	5
SIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
BRU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
PHI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
INS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	5

Table A2.
First-order contiguity matrix

Corresponding author
Seung-Whan Choi can be contacted at: whanchoi@uic.edu