

Metulini, Rodolfo

Conference Paper

Spatial gravity models for international trade: a panel analysis among OECD countries

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Metulini, Rodolfo (2013) : Spatial gravity models for international trade: a panel analysis among OECD countries, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123958>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Spatial gravity models for international trade: a panel analysis among OECD countries

Rodolfo Metulini

March 1, 2013

Abstract

The Gravity Model is the workhorse for empirical studies in International Economies and it is commonly used in explaining the trade flow between countries. Recently, several studies have showed the importance of taking into account the spatial effect. Spatial Econometric techniques meet this matter, proposing the specification of a set of models and estimators. We will make use of these Spatial Econometric techniques in order to estimate a Spatial Gravity of Trade for a 22-year-long panel of the OECD countries. The aim, therefore, is twofold: on one hand, we are going to use the newest Spatial Econometric techniques in a field where they aren't widely applied. On the other hand, we provide an updated interpretation of the behaviour of the International Trade in an OECD context, taking into account potential spatial spillover effect due to the third country dependence, and the effects of the migratory phenomenon.

1 Introduction

Literature shows how, at empirical level, the classical Gravity Model (Anderson 1979, Anderson, Van Wincoop 2003) brings to good results in explaining international trade. Anderson (1979, p.106) states that this equation is the most successful in explaining this issue, furthermore, Everett and Hutchinson (2002, p-489) define this model as the workhorse for empirical studies in international economy. Since the gravity model have physical roots, the trade flow depends on the dimension of both the origin and the destination country, as well as the distance between them.

Recently, several studies highlighted that an additional effect to take into consideration is the spatial dependence. Following Tobler first law of Geography (1970) -everything is related to everything else, but closer things are more closely related than distant things - the regional scientists began to give relevance to this aspect. In the International Trade, spatial dependence can be justified by the role of the *third country effect*.

The so called third country effect is connected with two different phenomena: the locational factor theory that fosters the spatial spillover effects, and the

persistence phenomena. For the first phenomena, it is asserted that, if some structural change happening in one country that force its trade flow, force the trade flow of the neighbors as well, producing positive effects on the volume of flows. On the other side, the third country can fosters persistence effect, which is based on the relative trade cost between countries i and j compared with the cost between i and k , where k is the third country. Assuming this phenomena, an increase on competitiveness of the third country lowers its trade cost and, consequently, lowers the flow between the couple ij .

From the beginning of the gravity model of trade, the space is taken into account using the distance variable and other variables considered relevant on the volume of trade, such as geographic contiguity, common language and common currency, and the presence of some Free Trade Agreements. However, these variables alone could not be satisfactory if the aim is to consider the dependence in space. The Econometric literature suggest to model this phenomena with appropriate spatial econometric models and methods.

In the Spatial Econometric field several models and estimation methods (see Anselin, 1988) have been proposed to meet this economic matter: Spatial Autoregressive Model (SAR) emerges when we view the spatial dependence as a long run equilibrium of an underlying spatio-temporal process. The Spatial Error Model (SEM) emerges when omitted variables exhibit spatial dependence, or in cases of economic shocks; however, as the choice of defining the spatial dependence on the error terms rather than on the dependent variable is a crucial point, the SARAR configuration, that merges both models, is also considered. The model's choice can follow two different approaches. Traditionally, Specific-to-General approach has been proposed: we start with the demeaned model and then add additional components. However, as discussed by Hendry (2004) and Florax, Folmer, Ray (2003) to mention a few, the model's choice could follow the General-to-Specific. In this case we start with the full model and then remove unnecessary components. Regarding the estimation methods, take into account the spatial dependence with auto regressive models which consist in considering the spatially lagged dependent variable in the set of the explanatory variables. As the dependent variable is likely to be correlated with his spatial lag, the endogeneity problem emerges, and ordinary least square (OLS) is not unbiased anymore. Several estimators are available: The traditional one is the Concentrated (or two stages) Maximum Likelihood proposed by Anselin (1988) and revised by LeSage, Pace (2008) for the flow models; a novel alternative is the Instrumental Variable Generalized Method of Moments (IV/GMM), that uses some specific procedure related to the choice of a set of appropriate instrumental variables jointly with the choice of appropriate moment conditions. This kind of estimation methods were developed by Kelejian, Prucha (1998) for sectional data, by Kelejian, Prucha, Kapoor (2007) and Mutl, Pfaffermayr (2008) for longitudinal data.

Another effect we want to consider is the migratory phenomena: recent literature on the relation between stock of immigrants and commercial flow (Gould 1994, Nijkamp et all 2011) suggests a positive relation. In fact, the stock of immigrants in the partner country increases commercial flows, since it

reduces the barriers represented by difference in culture, language and currency, and because the immigrants search for their home-country goods.

With this work we are going to estimate a spatial gravity model for the analysis of International Trade over a 22-years-long panel of the OECD countries. The aim, therefore, is twofold: by the first, we are going to apply the newest spatial econometric techniques in a research field in which this techniques are scarce. On the other hand, to provide an updated interpretation of the behavior related to international trade in an OECD context, accounting for the *spillover effect* and highlighting the persistence effect. Moreover, accounting for the migration phenomena theory. In Fact, although there are several panel gravity of trade analysis in the OECD context, using fixed effect rather than the random effect, and with the selection of different explanatory variables depending on the different goal of the analysis (Egger 2002, 2004, Rose 2000 to mention a few), the contributions that consider the spatial effect in this area are scarce.

This paper is organized as follows: in section 2 we present the motivations that underlie this work. In section 3 we show the economic and the spatial econometric aspects related to the gravity model for international trade. Section 4 presents the data and the results of the empirical analysis; finally, in chapter 5 we approach to conclusions and provide some arguments to debate about future contribution.

2 Motivations

Since its birth, the gravity has proved to be an excellent model for the empirical analysis of the effects that different variables have in relation to the international trade. Traditionally, as well as to measure the effect of distance and economic dimension, the gravity model has been used to assess trade policy implications and to analyze the effects of Free Trade Agreements.

Nonetheless, recently, several studies highlighted the importance of taking into account the effect of the spatial dependence in the gravity of trade as well as in the other international economics fields. The fast emergence of the first law of geography by Tobler (1970): *everything is related to everything else, but closer things are more closely related than distant things* led many researchers to analyze the spatial dependence by means of the gravity model formulation.

Such spatial dependence is due to the role of the *third country* (in other words, the role of the neighbours), which is gaining a central role in recent years to examine the dependence structure over space in international economics. Authors such as Baltagi (2007, 2008) and Hall, Petroulas (2008) show, at theoretical and empirical level, the importance to consider the effect of structural changes in a country on their spatial neighbours.

Traditionally, the *space* was taken into account using the variable *distance* and other variables considered relevant in the determining process of the trade flows (such as geographical contiguity, common language and common currency, and the presence of free trade agreements). Nevertheless, these variables may be inadequate since we assume that the behavior of each country has an ef-

fect on the behavior of their spatial neighbours. In fact, theories regarding the locational factors foster a phenomenon called *spatial spillover*. According to regional scientists, changes in locational factors in a country are related to changes in neighbouring countries, so, we can expect that, if these changes lead to an improvement in trade activities, for the spatial spillover effect, this will produce improvements in trade activities of the neighbouring countries as well.

The *third country* determines the spatial spillover, which is a positive effect. But, it can determine a negative effect too, that is a reduction of trade between couples. This is the persistence effect: the choice to trade from country i to country j depends on the relative cost (which is the cost from i to j compared with the cost from i to k , where k is the neighbour of j (i.e. the third country, Adam, Cobham, 2007) So, if the couple ij have near them a country that increases its competitiveness (greater demand for imports and exports, lower production costs, etc ...), then, the trade between this couple will decrease in favour of an increasing of the trade with the country k (Kelejian, Tavlas, Petroulas, 2011).

The persistence effect is analyzed in the context of the gravity equation by Kelejian, Tavlas and Petroulas (2011). In their analysis, the spatially lagged variables (the persistence effects) are represented by the GDP's of the origin countries (understood as a potential exporter competitor) and the GDP's of the destination countries (as a potential importer competitor), weighted by the inverse of the distance. The empirical analysis of these authors shows that the spatially lagged GDP from the export side has a not significant effect on trade, while the spatially lagged GDP from the import side has a negative effect. Such empirical evidence allows us to study the persistence effect.

Another factor we are interested in considering is the effect of migration on trade, in fact, recently, the theory about the relation between *stock* of immigrants and trade flows is undergoing a turnaround: Gould (1994) was one of the first to consider migratory phenomenon as an incentive element for international trade flows, in opposition to what was the point of view before, based on the theory on trade of *emphHescher – Ohlin*, for which international trade and migration are complementary, since countries choose between imports of goods or imports of labour. The reasons (at a *Macro* level) why the newest theories tend to consider the stock of immigrants as an incentive component on trade, is the fact that immigration increases aggregate demand which, in turn, increases the demand for imports too. Exports also increase, if immigration leads to a decrease on the unitary costs of production. In addition, the presence of immigrants can lead to a reduction of the effect of barriers to trade represented by the different language, different currency and cultural differences that may discourage the exchange between a pair of countries, bringing some specific *skills* and *knowledge* that foster the international trade. A final reason that can foster an increase in exports can be represented by the fact that immigrants tend to require origin-country products, which need to be imported.

A brief analysis of the literature conducted by Nijkamp et al. (2011) found that the majority of the empirical studies show the presence of this positive

effect, which allows us to model the *stock* of immigrants in the gravity model with the expectation of obtaining a positive sign of the estimated coefficient.

In the light of the facts mentioned above, our goals will be to grasp the potential contribution of spillover effects and migration data by using spatial gravity models and to produce an updated interpretation of the behaviour of International trade, since, despite a wide range of theoretical studies, the existing empirical evidence on the importance of the *third-country* and *migrational* effects in trade analysis is rather limited.

3 Economic and Econometric Review

The gravity model is the most used empirical model in international economics for its excellent empirical robustness in describing trade flows, although in its first years after its formulation by Tinbergen (1962), it lacked for a strong theoretical foundation. It describes trade flow between two countries as a function of the economic size of both the origin and the destination one, as well as of the geographical distance between each of them. The gravity model permits to analyze the main determinants of the trade flows, the presence of geographical effects, on one hand, and political effects, on the other, that limit the flow, and the existence of positive effects attributed to the Free-Trade Agreement. In most of the applications based on the original gravity model formulation by Tinbergen it is possible to split between two groups of variables, relatively to their effect, that could be *push* (positive contribution to the flow) or *pull* (negative contribution). The push factor is represented by the economic size and the gross domestic product (GDP) of the origin as well as the destination country is the main proxy. We can use other variables, such as population and GDP per-capita. The pull factor is represented by the transport costs of the traded goods, and it is possible to use the geographical distance between the two countries as an appropriate proxy. Other variables could be inserted in the model to specify the push effect, such as dummy variables representing if the two countries have common border, or if they share the same official language, or if they use the same currency; Rose (2000) was the first to introduce these variables in the gravity model. Furthermore, other factors have been considered, for example the free trade agreements, in order to identify a country as a part of some agreement, like the EU15, NAFTA or EFTA. The first theoretical - based gravity model derived by Anderson (1979) results from the equality of supply of export and import demand in the case of Constant Elasticity of Substitution (CES) expenditure function and Constant Elasticity of Trade (CET) profit equation¹ is generically represented by the equation 1.

$$Y_{ij} = \frac{X_i^{\beta^o} * E_j^{\beta^d}}{d_{ij}^{\theta}}. \quad (1)$$

¹it is furthermore based on the economic framework of Ricardian comparative costs and Heckscher-Ohlin factor endowments theories, Chamberlain monopolistic competition and increasing return to scale

This is the simplified form of the more inclusive model in equation 2, assuming that $P_j^{1-\sigma}$ and Π_i are equal.

$$Y_{ij} = \frac{X_i^{\beta_o} * E_j^{\beta_d}}{d_{ij}^{\theta} * P_j^{1-\sigma} * \Pi_i}. \quad (2)$$

Where Y_{ij} is the flow from country i to country j , G represents a constant factor, X_i and E_j represent the economic size respectively of the origin country and destination country; d_{ij} is the distance between the two countries (as a proxy of the transport cost), Π_i measures the market openness for the origin country (exporting goods to other countries) and P_j the market openness of the destination country (importing goods from other countries). Furthermore, β 's, and θ 's are the model coefficients to be estimated.

Baldwin, Taglioni (2006) dissert on the applicability of the model equation shown in 1. That model assumes that the coefficients measuring the openness of the origin and destination countries above defined, are equal, and they do not change over time. If this does not hold, it would be necessary to account for these coefficients, called balancing factors. As this assumption is not likely to be valid in panel data, it is demonstrated that we can solve the problem introducing a set of individual dummies for the origin countries and for the destination countries.

As we have discussed the economic specification of the gravity model in the preceding section, is now appropriate to present the econometric context in which this contextualizes. A wide variety of model had been specified in the spatial economic literature which can be adapted for the analysis of international trade with the use of the gravity model. The main models, proposed by Anselin (1988) for cross sectional specification are called SAR (spatial autoregressive model) and SEM (spatial error model). It is possible to derive each of those econometric models starting from an economic motivation. The motivation for the SAR model comes from viewing spatial dependence as a long-run equilibrium of an underlying spatio-temporal process; the motivation for the SEM model asserts that omitted variables that exhibit spatial dependence lead to a model with spatial lags of both the explanatory and the dependent variable. Referring to the first motivation, we can think to the spatial dependence as based on a time-lag relationship describing a diffusion process over space. SAR models contain spatial lags of the dependent variable. The second econometric motivation leads us to a model with a spatial lag of both the explanatory and the dependent variables. However, if the included and excluded explanatory variables are not correlated, a spatial error model (SEM) emerges.

Taking into account the spatial dependence (or, in other words, to represent in formula the spatial lags) consists on defining the structure of the spatial components, that involves the weights matrix. In a typical cross sectional model with n countries and one observation per country, the spatial weights matrix that combines neighbors, called W , is a n by n dimension, non negative and

sparse. For the elements of this matrix hold that $w_{ij} > 0$, and, by convention, $w_{ii} = 0$, to prevent an observation from being defined as a neighbor to itself. Given an origin-destination organization of the data, as flow data, we should define a weight matrix of order n^2 by n^2 that combines each couple of regions having neighboring origin country or neighboring destination country. Given the above defined W , we can define the n^2 by n^2 matrix W_{od} as the kroeneker product of W with itself, which keep into account the interaction between origin and destination neighbors. We can now define the SAR (equation 3) and the SEM (equation 4) model derivation of the gravity equation

$$Y = I_n \otimes I_t \otimes \alpha_i + I_n \otimes I_t \otimes \alpha_j + I_n \otimes I_n \otimes \alpha_t + X_o \beta_o + X_d \beta_d + D\gamma + \rho(W_{od} \otimes I_t)Y + \varepsilon \quad (3)$$

$$Y = I_n \otimes I_t \otimes \alpha_i + I_n \otimes I_t \otimes \alpha_j + I_n \otimes I_n \otimes \alpha_t + X_o \beta_o + X_d \beta_d + D\gamma + \lambda(W_{od} \otimes I_t)\varepsilon + u. \quad (4)$$

Where Y is a vector of dimension $n*n*t$, I_T is a vector of ones of dimension T and I_n is a vector of ones of dimension n . α_i and α_j are the vector of order n that represent the origin and destination fixed individual effect, and α_t is the T dimension time counterpart. Moreover, X_o and X_d are the matrices of order $n*n*t$ by k , that contain respectively the k explicative variables of the origin and destination countries. β_o and β_d are the k by 1 vector of coefficients for the explicative variables. γ is the scalar coefficients relative to D , that is the $n*n*t$ by h matrix with the h variables (like distance) relative to the interaction between the origin and destination countries. Moreover, ε in equation 3 is the $n*n*t$ vector of the residuals, that is assumed to be a stochastic i.i.d. variable with zero mean and common variance σ_ε^2 . ρ is the coefficient for the spatial dependence captured by the spatially lagged dependent variable, W_{od} is defined above and I_t is a t by t identity matrix that account for the time index. The term u in equation 4 is i.i.d with zero mean and σ_u^2 variance. Here, we assume a simplification of the model in 5, derived from filtering the dependent variable from spatial components, that would leads to the following model (as defined in LeSage, Pace, 2008).

$$Y = I_n \otimes I_t \otimes \alpha_i + I_n \otimes I_t \otimes \alpha_j + I_n \otimes I_n \otimes \alpha_t + X_o \beta_o + X_d \beta_d + D\gamma + \rho_1(W_o \otimes I_t)Y + \rho_2(W_d \otimes I_t) + \rho(W_{od} \otimes I_t)Y + \varepsilon \quad (5)$$

The economic motivations that leads to the two different econometric models could be, anyway, complementary, so, we can jointly take these into account. The model that in this case emerges is a SARAR model with spatial lags of both kinds.

$$Y = I_n \otimes I_t \otimes \alpha_i + I_n \otimes I_t \otimes \alpha_j + I_n \otimes I_n \otimes \alpha_t + X_o \beta_o + X_d \beta_d + D\gamma + \rho(W_{od} \otimes I_t)Y + \lambda(W_{od} \otimes I_t)Y\varepsilon + u. \quad (6)$$

Defined the model, we have to choose the best estimation method to end up with correct coefficients. Every spatial econometric model suffers from endogeneity problems, since the spatially lagged component is naturally correlated with the dependent variable itself. Since Anselin (1988), the more common estimation method to overcome the problem of the endogeneity is the concentrated maximum likelihood. A novel approach based on the methods of moments and the instrumental variable (IV/GMM) was formulated later by Kelejian, Prucha for the Cross section models (1998), Kapoor, Kelejian, Prucha (2007) and Mutl, Pfaffermayr (2008) (to mention the most representative works) for longitudinal data. Maximum likelihood estimation involves maximization of the concentrated log-likelihood function (concentrated for σ^2 and the coefficients of the explanatory variables $(\beta_o, \beta_d, \gamma)$ with respect to the spatial autoregressive parameters ρ and λ . Although? this is the workhorse estimator in spatial econometric models, a big difficulty appears in the numerical maximization of the concentrated log-likelihood function on evaluating the n by n log-determinant of $A = I_{n^2} - \rho W_{od}$ and $B = I_{n^2} - \lambda W_{od}$, that are the factors to filter out the spatial effect in order to solve the concentrate log-likelihood. A solution could be to approximate the weights matrix with a sparse matrix use, using the sparse matrix Cholesky factorization techniques. The GMM accounts for these difficulties and it is feasible for large sample size. Results on efficiency by Das, Kelejian, Prucha (2003) and Lee (2003) show that both the GMM and the IV estimators are as efficient as the corresponding ML estimators in small samples, but overcome the computational problems deriving from evaluating the log-determinant of the determinant of the matrix for large sample. The method of Generalized method of moments was for the first time proposed by Kelejian, Prucha (1998) for cross sectional data and not accounting for spatial lag of the dependent variable. Later, Kelejian, Prucha, Kapoor (2007) implement the same procedure for a panel data with random effect; Mutl, Pfaffermayr (2008) proposed the methods for a panel data, that compound the IV techniques with the GMM for random as well as fixed effect models.

4 Application

4.1 Data, models and methods

We want to model the country-to-country trade flow over the period from 1988 to 2009 for the OECD members. The sample was restricted to the 32 OECD countries² for $n=22$ years, resulting in a $n*n*T = 22528$ observations. Exportations, deflated to obtain a value on constant prices, are taken into consideration. As we deflate through the CPI (consumer price index) relative to the United States with year base 2000, a consideration emerges: Baldwin and Taglioni (2006), in what they called bronze medal error, asserted in fact that this choice brings to a biasness in the estimation. We correct this bias by putting the time effect into the model (as proposed by Rose, 2000). One other trick is to intro-

²See list in table 1

Table 1: List of OECD countries*

Austria	Japan
Australia	Mexico
Belgium-Luxembourg	Netherlands
Canada	New Zealand
Czechoslovakia	Norway
Chile	Poland
Denmark	Portugal
Estonia	South Korea
Finland	Israel
France	Slovenia
Germany	Spain
Greece	Sveden
Hungary	Switzerland
Iceland	Turkey
Ireland	Great Britain
Italy	United States

(*) We merged Belgium and Luxembourg for the 1999-2009 interval, and Slovakia and Czech Republic for the 1994-2009 interval.

duce the purchase power parity (*ppp*) index to account for the global time-trend of the US CPI. The exportations data are available at the web site of UnComTrade. UnComTrade make available this variable in nominal data (in terms of US dollars). The CPI of the US at base year 2000, is, instead, available at OECD web site.

There is a global trend in the inflation rate when we deflate the nominal value through the consumer price index (*cpi*) of the United States. It is possible to fix this issue by including the consumer prices index variable expressed in purchasing power of parity (*ppp*), which represents the cost of goods in the basket of a country, divided by the cost of goods in the U.S. basket. In the long term, in fact, the exchange rate and the *cpi* tend to adjust to the value of the consumer price index in purchasing power of parity (in 3-4 years).

So, if the right model includes as a dependent variable $\frac{Export}{ppp}$ and we use something like $\frac{Export}{cpi}$ ³, we insert between the explanatory, the term $\frac{ppp}{cpi}$ ⁴

The explanatory variables of the model are the index of the purchase power of parity (*ppp*), the gross domestic product (*gdp*) in real terms and the population (*pop*). Moreover, besides the dummies *contig*, *comcur*, and *comlang*, the dummies relatives to the free trade agreement of *EU15*, *NAFTA* and *EFTA* take part into the model. Furthermore, the dummy variables representing the individual effect of both origin and the destination countries are inserted in the

³*Export* is identified with the dependent variable, represented by exports in nominal values

⁴the reason is that the *cpi* has a global trend: it is affected by other components that cause spurious correlation (short-term volatility due to external factors)

model.

The purchase power parity index is available at PennTable web site, in a dataset called PWT 7.0. The explanatory variable, as well, are available at the database of PennTable (*gdp* and *pop*), and from CEPII (*comlang*, *contig* and *comcur*).

We have, furthermore, economic reasons to introduce a set of individual effects that account for the balancing factor and a set of time effects accounting for global trends. Although it is recognized the preference of the fixed effects versus random effects in the gravity model applications to trade flows, it is necessary to point out the differences between the two approaches. Since with the fixed effects, we have n dummies for the origin country, n dummies for the destination countries and t dummies for the years, the number of additional parameters to be estimated are $2n + t$. A natural consequence is the need of a large panel with large n as well large t to avoid the problem of identification. With the random effect the assumption of independency between effects and the residuals need to be hold, however, this empirically hardly happens in gravity of trade; the fixed effects do not require this assumption, so it leads us to favor the fixed effects. Furthermore, recent works (Elhorst 2010) show how the use of random effects is justifiable in the case of samples data, while, if we work with the entire population (which is generally, the case of works in international trade flows), fixed effects are preferred.

Along with the spatial effects components, we have defined two different matrices, the first with the contiguity approach (it was used *spmat* command in STATA and the coordinates of the countries availables at CEPII), the second one using the inverse of the distance between the centroids (available at CEPII as well). We have used Moran I index (calculated for every year over the period from 1988 to 2009) to discriminate between the better one. The results of these tests are presented in table 2 of the appendix, and show an higher value of the index for the contiguity-based matrix, although for both the matrices the index is significant. This strong difference between the two indices should be due to the fact that the inverse distance matrix connects, all the spatial units, albeit with different weights. Therefore, although the Moran index is lower, the inverse distance matrix represents a more comprehensive vision of spatial relations. For this reason, we consider the inverse distance matrixes the best one. As proposed by Arbia (2009) we check the spatial effect assumption by means of the Moran I on the OLS residuals with the inverse distance matrix. This test shows a significant spatial effect, as we can see on the second and the third column of table 3.

As we are allowed to do so, we can introduce the spatial component in the model, we assume, by adopting a *General-to-Specific* approach, that the true model is the SAC in which the spatial lag of both the dependent and the error is taken into account. In order to choose between random versus fixed effect specification, a spatial Hausmann test was performed. This test highlights the strong preference for the fixed effect model (see the first column of table 5 in the appendix). Furthermore, a *SpatialLagrangeMultiplier* test was performed to

Table 2: *Moran-I* of the inverse distance matrix*

YEAR	Inverse Distance	P-Value	Contig	P-Value
1988	0.083	$< 2.2e - 16$	0.436	$< 2.2e - 16$
...
2009	0.106	$< 2.2e - 16$	0.470	$< 2.2e - 16$

(*) the *Moran - I* for every year in the middle is significant.

Table 3: *Moran-I* on regression residuals to test the spatial effect*

YEAR	OLS residuals	P-VALUE	SAR residuals	P-VALUE
1988	0.198	$< 2.2e - 16$	0.117	$< 2.2e - 16$
...
2009	0.153	$< 2.2e - 16$	0.054	$< 2.2e - 16$

(*) These results are obtained using the W_{od} constructed with the inverse of distances.

The *Moran - I* for every year in the middle is significant.

check the presence of spatial effects and spatial autocorrelation out. Each of the three tests used for this purpose is significant (test for the presence of spatial autocorrelation assumed the presence of spatial effects, tests for the presence of spatial effects assumed the presence of spatial autocorrelation, join test), as we can see in the column 2-4 of table 4. These tests would lead to a SEM model, but the significance of additional Moran I indexes calculated on the SEM residuals (columns 4-5 of table 4) leads us to consider as true the SARAR specification.

We have estimated the IV/GMM with *spgm* command in *splm* package using *within* model and fixing the spatial effect in both the error and the dependent variable as *true*.

To sum up, the empirical model takes the following form:

$$\begin{aligned} \text{Export}_{ijt} = & \alpha_{ij} + \alpha_t + \beta_1^o \text{Pop}_{it}^o + \beta_2^o \text{GDP}_{it}^o + \beta_3^o \text{ppp}_{it}^o + \beta_4^o \text{Nafta}_{it}^o + \beta_5^o \text{Efta}_{it}^o + \\ & \beta_6^o \text{Eu15}_{it}^o + \beta_1^d \text{Pop}_{jt}^d + \beta_2^d \text{GDP}_{jt}^d + \beta_3^d \text{ppp}_{jt}^d + \beta_4^d \text{Nafta}_{jt}^d + \beta_5^d \text{Efta}_{jt}^d + \beta_6^d \text{Eu15}_{jt}^d + \\ & \beta_7 \text{Migrat}_{ijt} + \psi_1^{od} \text{Contig}_{ij} + \psi_2^{od} \text{Comlang}_{ij} + \psi_3^{od} \text{Comcur}_{ij} + \psi_4^{od} \text{Dist}_{ij} + \\ & \psi_5 w_{ij,hk,OD} \text{bilat.PIL}_{ijt} + \rho w_{ij,hk,OD} y_{hk,t} + u_{ijt}; \end{aligned}$$

$$u_{ijt} = \lambda \sum_{hk=1}^{n^2} w_{ij,hk} u_{hk,t} + \varepsilon_{ijt}; \quad (7)$$

Table 4: Diagnostic tests for the SARAR model specification*

Hausmann test		LM-H tests**		LM- λ tests***		LM- μ tests****	
chisq	p value	value	p value	value	p value	value	p value
11021.45	< 0.01	112687.8	< 0.01	80.85	< 0.01	313.69	< 0.01

(*) I test del Lagrange Multyplier sono stati svolti anche su specificazione SAR e SEM, fornendo valori in confronto piu' bassi.

(**) Baltagi, Song, Koh *LM Conjoint test* for spatial autocorrelation and random effects.

(***) Baltagi, Song, Koh *LM conditioned test* for spatial autocorrelation (assuming random effects > 0).

(****) Baltagi, Song, Koh *LM conditioned test* for random effects (assuming spatial autocorrelation > 0).

where:

- $Export_{ij,t}$ are the real values exportations of *reporter* and *partner* countries;
- Pop_{it}^o e Pop_{jt}^d are the population of *reporter* and *partner* countries;
- ppp_{it}^o e ppp_{jt}^d are the consumer price indexes at purchase power of parity of *reporter* and *partner* countries;
- GDP_{it}^o e GDP_{jt}^d represent the gross domestic products at real values of *reporter* and *partner* countries;
- $Nafta_i^o$ e $Nafta_j^d$ are a set of *dummies* to identify the belongs to the *Nafta* free trade agreement for *reporter* and *partner* countries;
- $Efta_i^o$ e $Efta_j^d$ are a set of *dummies* to identify the belongs to the *Efta* free trade agreement for *reporter* and *partner* countries;
- $Eu15_i^o$ e $Eu15_j^d$ are a set of *dummies* to identify the belongs to the *EU15* free trade agreement for *reporter* and *partner* countries;
- $Comlang_{ij}$ are a set of *dummies* which identify if the couple have common borders;
- $Comcur_{ij}$ are a set of *dummies* which identify if the couple have the same currency;
- $Contig_{ij}$ are a set of *dummies* which identify if the couple speak the same language;
- $Dist_{ij}$ represents the geographical distance between the couple of countries;

Table 5: *IV/GMM* results for SARAR model

VARIABLE	Estimate	Std.Error	T-values	Pr(> t)
W Bilat. PIL	-0.0802	3.4755e+01	-23.0633	< 2.2e-16 ***
Migrat	0.0321	4.2999e-03	-7.4643	8.377e-14 ***
Contig	0.3711	2.9953e-02	12.3890	< 2.2e-16 ***
Comlang	1.0478	2.8075e-02	37.3198	< 2.2e-16 ***
Dist	-0.6990	7.8894e-03	-88.5944	< 2.2e-16 ***
Comcur	2.4469	1.2805e-01	19.1085	< 2.2e-16 ***
Pop ^d	1.5369	2.8829e-02	53.3112	< 2.2e-16 ***
ppp ^d	0.3057	1.2207e-02	25.0430	< 2.2e-16 ***
PIL ^d	2.5915	6.6409e-02	39.0228	< 2.2e-16 ***
Pop ^o	1.8068	4.3787e-02	41.2643	< 2.2e-16 ***
ppp ^o	0.4639	1.6308e-02	28.4498	< 2.2e-16 ***
PIL ^o	2.3860	6.5350e-02	36.5115	< 2.2e-16 ***
Efta ^o	1.9269	7.1639e-02	26.8975	< 2.2e-16 ***
Efta ^d	0.3981	3.0510e-02	13.0480	< 2.2e-16 ***
Nafta ^o	-2.5332	1.1428e-01	-22.1660	< 2.2e-16 ***
Nafta ^d	-1.9925	7.6356e-02	-26.0943	< 2.2e-16 ***
EU15 ^o	0.7937	2.2052e-02	35.9905	< 2.2e-16 ***
EU15 ^d	0.7236	1.8046e-02	33.7602	< 2.2e-16 ***
Rho	0.00035	-	-	-
Lambda	0.00230	3.2513e+02	-23.7515	< 2.2e-16 ***

- $Migrat_{ijt}$ represents the *Stock* of immigrants from country i to country i .
- $Bilat.PIL_{ijt}$ is the weighted mean of the *GDP's* of *reporter* and *partner* countries;⁵

4.2 Results

As confirmed by the findings of the empirical literature, the use of the gravity model in explaining the trade flow phenomena between OECD countries leads to a good result in terms of fitting. Table 5 shows the results of the estimated *IV/GMM* for the SARAR model. Observing the results of the estimated *IV/GMM* coefficients in table 5, we can observe that the signs of the structural components of the model are similar to what literature shows.

The gross domestic product of the *reporter* (GDP^o) and *partner* (GDP^d) countries are positive, as expected. The population (Pop^o and Pop^d) plays an important role, with a positive and significant sign for both the estimated coefficients. It appears that the distance has a negative effect on trade, with

⁵ $Bilat.PIL_{ijt} = \frac{PIL_{it}^o + PIL_{jt}^d}{2}, \forall i = 1, \dots, n; \forall j = 1, \dots, n$

an estimated value of the coefficient that reflects the literature. As expected, common language (*Comlang*), common borders (*Contig*) and same currency (*Comcur*) play a positive role in the increasing of international volume of trade. This proves that the discriminatory policies based on cultural and economic factors are still important in determining the trade flow, as shown by Bang (2006). The free trade agreements are also significant: belonging to the European Union ($EU15^o$ and $EU15^d$) and to *EFTA* ($EFTA^o$ and $EFTA^d$) increases the flow of trade. The interpretation that emerges for agreements between the countries of *NAFTA* ($Nafta^o$ and $Nafta^d$), however, leads to an ambiguous considerations : in fact, the estimated *IV/GMM* coefficients shows a negative sign. This occurrence, however, could be due by the fact that the countries taking part of *NAFTA* are *USA* and *Mexico*, which are highly populated. Since these could be considered as outliers, the model may estimate too much higher coefficients of population and then may correct this coefficient by the *NAFTA* dummies, that, consequently, are not going to capture only the effects of free agreement, but also to correct the outliers effect described above.

In support of the theory on persistence effect, the *GDP* of the neighboring countries ($W\ Bilat.*GDP$) has a negative coefficient. Instead on the contrary?, in support of the new theories on the effect of the *stock* of immigrants, the variable *Migrat* has a positive effect.

The last two rows of Table 5 show the estimated value of the spatial coefficients, ρ for the effect of the *autoregressivemodel* and λ for the effect of the *error model*. Albeit with quite low absolute value, the corresponding positive and significant signs seem to confirm the role of spatial dependence in the analysis of international trade among *OECD* countries.

5 Conclusions and further developments

In this analysis we studied the pattern of trade among the *OECD* countries, using the gravity model. We have also formalized the model, both from an economic point of view and from a econometric side, giving emphasis to the spatial aspect. We used the main and newest spatial econometric techniques for the specification of the spatial component. The results show a good fit of the gravity model to the data relating to the trade among *OECD* countries, confirming the importance of structural variables of the theoretical model and an empirical confirmation of the theories that motivated this analysis: the presence of spatial dependence that motivates the choice of a *SARAR* model with both lagged dependent variable and lagged error terms. This brings to the confirmation of the theory relating to *spillovers* and *locational factors*. The new theories about the effect of migration and theory on the persistence effects are further supported.

However, there are some points for which the analysis deserves further consideration, both from the economic point of view, and from the econometric modeling. Although the methods based on spatial autoregressive models are

widely discussed and considered in the field of econometrics, the application of these methodologies adapted on gravity equation has not been widely discussed yet. The model for flow data proposed by Le Sage and Pace (2008), which considers separately the spatial dependence between *reporter* countries and between *partner* countries, estimated using a concentrated maximum likelihood method, could be extended to the longitudinal context. In addition, these models could be estimated by a *IV/GMM* approach.

Another aspect of research can be related to the introduction of the dynamic component in a *SARAR* model. several models and estimation methods have been recently proposed in the context of *panel* data, (Baltagi et al., 2011, Jacobs et al., 2011). In this case, the proposed estimator for dynamic models (*SystemGMM*) is based on first differences data transformation, mixed with the *IV/GMM* estimator, using *Within* transformation before.

Finally, at a level of economic interpretation, it could be interesting to investigate if the *global – trend* can be captured by some specific variables related to the financial openness, innovative activities and environmental policies (Costantini, Mazzanti, 2010).

6 Bibliografia

- Adam C., Cobham D. (2007): Modelling multilateral trade resistance in a gravity model with exchange rate regimes. Centre for dynamic macroeconomic analysis conference papers 2007.
- Anderson J.A. (1979), A Theoretical Foundation for the Gravity Model, The American Economic Review, 69,106-116.
- Anderson J.A., Van Wincoop E. (2003), Gravity with Gravitas: A solution to the Border Puzzle American Economic Review 93, 170-192.
- Anselin L. (1988), Spatial econometrics: Methods and Models (Kluwer Academic Publishers, Dordrecht).
- Arbia, G. (2009), Spatial Econometrics: methods and applications Psycha Verlag (A Springer company).
- Baldwin R., Taglioni D. (2006) Gravity for dummies and dummies for gravity equations.- Working Paper 12516.
- Baltagi B.H., Song S.H., Koh W. (2003): Testing panel data regression models with spatial error correlation. Journal of econometrics 117, 123-150.
- Baltagi BH (2003), A Generalized design for bilateral trade flow models, economic letters - Elsevier.
- Baltagi BH et al. (2007), Testing for serial correlation, spatial autocorrelation and random effects using panel data, Journal of Econometrics.

- Baltagi B.H., Egger P., Pfaffermayr M. (2007): Estimating models of complex FDI: Are there third-country effects. *Journal of Econometrics*, Vol. 140, Iss. 1, Pp. 260-281.
- Baltagi B.H., Egger P., Pfaffermayr M. (2008): Estimating regional trade agreement effects on FDI in an interdependent world. *Journal of Econometrics*, Vol. 145, Iss. 12, Pp. 194-208.
- Baltagi BH (2008), *Econometric analysis of Panel Data*, 4th edition, John Wiley e sons. Ltd.
- Baltagi B.H. (2011) - Estimating and Forecasting with a Dynamic Spatial Panel Data Model - SERC discussion paper.
- Bang J.T. (2006): Regional Integration, Spatial Effects, and the Gravity Equation.
- Cizek P., Jacobs J.P.A.M., Ligthart J. E., Vrijburg H. (2011): GMM Estimation of Fixed Effects Dynamic Panel Data Models with Spatial Lag and Spatial Errors. Center Discussion Paper No. 2011-134.
- Costantini V., Mazzanti M. (2010): On the Green Side of Trade Competitiveness? Environmental Policies and Innovation in the EU. FEEM Working Paper No. 94.2010.
- Das D., Kelejian H.H., Prucha I. R. (2003) - Finite sample properties of estimators of spatial autoregressive models with autoregressive disturbances, *Papers Reg. Sci.* 82, 126 (2003).
- Egger P. (2000): A note on the proper Econometric specification of the gravity equation. *Economics Letters* 66, 25-31.
- Egger P. (2002): An Econometric View on the Estimation of Gravity Models and the Calculation of Trade Potentials. *The World Economy*, Vol. 25, Iss. 2, pp. 297-312.
- Egger P., Pfaffermayr M. (2004), The impact of bilateral investment treaties on foreign direct investment, *Journal of Comparative Economics*, Volume 32, Issue 4, Pages 788-804, ISSN 0147-5967, 10.1016/j.jce.2004.07.001.
- Elhorst J.P. (2010): Spatial Panel Models. In, *Handbook of Applied Spatial Analysis. Part 3*, 377-407. Springer.
- Everett, S.J., Hutchinson, W.K. (2002), The Gravity equation in international economics *Scottish Journal of political economy*, vol.49, no.5.
- Florax, J, Folmer H. Ray S. (2003) Specification searches in spatial econometrics: the relevance of Hendry's methodology, *Regional Sciences and Urban Economics* Vol. 33, Issue 5.

- Getis, A. (1995) Spatial Filtering in a regression framework: Example using data on urban crime, regional inequality, and government expenditure in : Anselin , L. , Florax, R.J.G.M. (eds), *New Direction in spatial econometrics*, Springer, Heidelberg, pp.172-185.
- Getis A. (1990) Screening for spatial dependence in regression Analysis, *Papers in Regional Sciences*, Vol, 69 Issue 1, pp. 69-81.
- Griffith D.A. (2000) Eigenfunction properties and approximations of selected incidence matrices employed in spatial analyses, *Linear Algebra and its Applications*, Vol. 321, Issue 1-3, pp.95-112.
- Gould M.D. (1994): Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows. *The Review of Economics and Statistics* Vol. 76, No. 2, pp. 302-316.
- Griffith , D.A. (2000), A linear regression solution to the spatial autocorrelation problem *Journal of Geographical Systems*, 2,141-156.
- Hall S.G., Petroulas P. (2008): Spatial interdependencies of FDI locations: A lessening of the tyranny of distance?. Dept. of Economics, University of Leicester.
- Kapoor M., Kelejian H.H., Prucha I.R. (2007): Panel Data Models with Spatially Correlated Error Components. Centre for analytical Finance Research Paper No.02-05.
- Kelejian H.H., Prucha I.R. (1998): a Generalized Spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of real estate finance and economics*, Vol.17:1,99-121.
- Kelejian H.H., Tavlak G.S. , Petroulas P. (2011): In the neighborhood: the trade effects of the euro in a spatial framework. *Regional Science and Urban Economics*, Vol. 42, Iss. 12, Pp. 314-322.
- Kepaptsoglou K., Karlaftis M.G., Tsamboulas D. (2010) - The Gravity Model Specification for Modeling International Trade Flows and Free Trade Agreement Effects: A 10-Year Review of Empirical Studies *The open economic journal*.
- Lee, L.F.(2003) - Best Spatial TwoStage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances, *Econometric Review*, Vol. 22, Issue 4. Elsevier.
- LeSage J. and Pace R.K., (2008). Spatial econometrics modeling of origin-destination flows. *Journal of Regional Science*, 48(5): 941-967.
- Mutl J., Pfaffermayr M. (2008): The Hausmann Test in a Cliff and Ord Panel Model. *Econometrics Journal*, volume 10, pp.1-30.

- Nijkamp, P. et al. (2011): Migrants and International Economic Linkages: A Meta-Overview, Tinbergen Institute Discussion Paper 11-147.
- Rose, A.K., (2000). One money, one market: Estimating the effect of common currencies on trade. *Economic Policy* 20 (April), 745.
- Tinbergen J. (1962), *Shaping the World Economy; Suggestions for an International Economic Policy* Twenty Century Fund, New York.
- Tobler W.R. (1970): A computer movie simulating urban growth in the Detroit region. *Economic Geography*, vol. 46.