QUADERNI DEL DIPARTIMENTO DI SCIENZE DELL' ECONOMIA

UNIVERSITÀ DEL SALENTO

A Gravity Model for Migration with Spatial Effect

Giada Andrea Prete

Numero E/19
Agosto 2020

DIPARTIMENTO DI SCIENZE DELL'ECONOMIA
FACOLTÀ DI ECONOMIA
UNIVERSITÀ DEL SALENTO - LECCE
ISSN 2284-0818

Quaderni del Dipartimento di Scienze dell'Economia

I Quaderni intendono favorire la circolazione di studi in versione preliminare, al fine di accelerarne la diffusione e per sollecitare suggerimenti e critiche.

Vengono considerati per la pubblicazione sui Quaderni del Dipartimento solo lavori non ancora pubblicati in versione definitiva su libri o su rivista. Possono proporre lavori per la pubblicazione membri del Dipartimento (anche in collaborazione con studiosi esterni) o studiosi che abbiano presentato il lavoro in oggetto nell'ambito di conferenze o seminari svolti nell'ambito delle attività dipartimentali.

I quaderni si articolano in quattro sezioni: Aziendale, Economica, Giuridica, Matematico Statistica.
I quaderni vengono regolarmente inviati a Istituti, Dipartimenti e Biblioteche a carattere economico e sono inclusi nella serie Research Papers in Economics (REPEC).

L'accesso ai Quaderni è approvato dal Coordinatore scientifico della sezione.

Editore:

Dipartimento di Scienze dell'Economia
Indirizzo mail: direzione.dipeconomia@unisalento.it
Ecotekne - via Monteroni
73100 Lecce

Codice ISSN: 2284-0818

A Gravity Model for Migration with Spatial Effects

Giada Andrea Prete*

Abstract

The aim of this paper is to suggest the estimation of a migration gravity model taking into account the presence of spatial interdependence between country-pairs. Considering as baseline the Random Utility Maximization Model, we discuss two fundamental issues that arise in estimating gravity equation in its log-linearized form through OLS: the presence of heteroskedasticity and the existence of zeros in migration data. Following Santos Silva and Tenreyro (2006), we propose the estimation of migration gravity models by using a Poisson Pseudo-Maximum Likelihood estimator (PPML) which is able to handle the mentioned issues and consequently leads to consistent estimates. Moreover, in order to capture spatial interdependence, we introduce a directed contagion effect, by using the approach suggested in the work of Kapetanios *et al.* (2014), which endogenously determines the neighbourhood or similarity between units according to a specific concept of distance through a threshold regime selection mechanism. In the empirical application, we find evidence of the presence of a diffusion effect in migration flows between different country-pairs in the context of European Neighbourhood Policy (ENP).

Keywords: Gravity models; Migration; Poisson Pseudo-Maximum Likelihood; Spatial effects.

JEL Classification: C33; F22.

1. Introduction

A lot of socio-economic phenomena have been explained by gravity equations. The instrument of Newton's law of gravity has been borrowed from physics and extended to investigate empirically the determinants of the movement of goods and services, capital and people. A general formulation of the gravity model assumes:

*University of Salento, Department of Economic Sciences, e-mail: giadaandrea.prete@unisalento.it

$$T_{ij} = \frac{f(R_{ij}, A_{ij})}{f(D_{ij})}$$

where the bilateral flows between country i and country j, T_{ij}, are a function of some repulsive or push factors, R_{ij}, and of attractive or pull factors, A_{ij}, given by the size of the economy of the involved countries, and an inverse function of the distance between origin and destination countries, D_{ij} (Patuelli *et al.* 2013). According to this specification, the determinants of migration can be distinguished into two big sets: push factors refer to all those elements which drive people away from their origin place, while pull factors represent a complex of attractive conditions which generate a favourable environment in the destination country. The function at the denominator corresponds to the costs that people have to deal with when they decide to move.

The attention on migration patterns has been addressed first by the seminal works of Ravenstein (1985, 1989) who traced some "laws of migration" which appear to guide all migratory movements. By looking at European and American migration, he observed that important determinants in people movements are due to the desire of human being to progress in social and economic conditions, to reach a more remunerative occupation, but this need is driven by the immediate neighbourhood of the destination place. Therefore, distance plays a crucial role in determining the mass of people moving from country to country.

For sure, thanks to the contribution of Tinbergen (1962) a gravity specification has been used to explain international trade flows. After him, Anderson (1979) presented a first theoretical justification for the gravity model based on constant elasticity of substitution (CES) preferences and goods that are differentiated by region of origin. Bergstrand (1989) and Deardoff (1998) preserved this structure of preferences and reconsidered the role of technology in terms of monopolistic competition and Hecksher-Ohlin structure. The two major contributions for theoretical microfoundations on gravity models are due to Anderson and Van Wincoop (2003,2004) and Beine *et al.* (2016). In the framework of general equilibrium gravity model, Anderson and Van Wincoop (2003,2004) developed a method which gives efficient and consistent estimates of gravity paying particular attention on trade barriers, and applied the theoretical model to solve the McCallum's border puzzle problem. In particular, this "border effect" refers to a situation in which there is higher volume of trade within a country compared with the volume of trade across the country's borders. On the other hand, Beine *et al.* (2016) provide an overview of the estimation of gravity models through Random Utility

Maximization models, which is able to describe the location decision problem that individual face to derive the expected probability to migrate according to their utility function. Both, Anderson and Van Wincoop (2003) and Beine *et al.* (2016) noticed that the specification suggested by the economic literature presents some drawbacks. In particular, these models ignore the fact that agents could be heterogeneous with respect to the perceived cost of migration and could be sensitive to variations in the attractiveness of alternative destinations. So that, in order to solve this trouble, the specification of the gravity equation should be rearranged in order to account for multilateral resistance terms.

The multilateral resistance (MLR) problem is defined as a form of bilateral trade barrier relative to average trade barriers that the pair of countries face with all the other partners. Following Ramos (2017), the concept of MLR is "related to the influence of third countries in determining migration flows between two particular countries". The traditional approach to bring out the multilateral resistance (MLR) effect is to take into account some fixed effects, in the specification of the gravity equation.

Ortega and Peri (2013) tried to control for MLR by including in the gravity model specification origin-year fixed effects, while Beine and Parsons (2015) include destination-year fixed effects. As explained by Ramos (2017) other more complex contributions have been proposed in literature. In the specification of Beine *et al.* (2016), the authors consider the presence of a time-invariant fixed effects for country pairs, so called dyadic effects and time effects to account for common unobserved shocks. The dyadic term gives the same contribution as time invariant variables that expresses a bilateral link, such as distance or dummies about common language or common borders.

In considering a more complex structure of the gravity equation model, it's important to evaluate the work of Bertoli and Moraga (2013) who introduced the cross-sectional averages of the dependent variable and of the regressors combined with dyadic fixed effects. This approach was defined with the Common Correlated Effects (CCE) estimator developed by Pesaran (2006).

In the same field, Serlenga and Shin (2007,2013) modeled the presence of multilateral resistance and cross-sectional dependence by referring to the studies of Pesaran (2006) and Bai (2009). Moreover, Behrens *et al.* (2012) consider the cross-sectional correlation between trade flows and error terms implemented in a spatial econometric framework.

This tendency of the existing literature to deal with the presence of unobserved and time-varying multilateral resistance and bilateral heterogeneity, controlling in some cases for cross sectional dependence, has been carried on by Mastromarco *et al.* (2016). In particular, they conjugate the factor-based approach adopted by

Serlenga and Shin (2013) and the spatial-based techniques used in Behrens *et al.* (2012) in order to determine the prominent role of Euro's trade effects. The study reveals the importance to take into consideration the historical and multilateral perspective of bilateral trade flows, without focusing only on the isolated shock induced by the monetary union. Additionally, Gunella *et al.* (2015) evaluated the Euro impact on trade on a dataset over 1960-2008 for 91 country-pairs of 14 EU countries by using observed and unobserved factors which were allowed to be cross-sectionally correlated in order to control for time-varying MLR.

In light of the above, the spatial techniques have a great relevance in disclosing forms of interdependence between countries in a dyadic setting. As considered in Neumayer and Plümper (2010) "what one unit does in relation to the other units, with which it forms a dyad, will often influence and be influenced by the relations of other dyads". In their work, Neumayer and Plümper tried to categorize different types of spatial effect modeling in both directed and undirected dyadic data. These forms of contagion can be combined with different specifications of the weighting matrix. When we talk about directed dyadic data we are referring to the interaction between two individual units forming a pair. The interaction between the units initiates with unit i (source country) and is directed towards j (target or destination country). It means that externalities could come from dyads which are different from the specific observed pair or by splitting the dyad in its two parts, *i.e.* the contagion effect could also rise from other sources or other targets. Generally, these forms of spatial interdependence are included in the model specification using a spatial lag of the dependent variable, which concerns the volume of migration flows between country-pairs. In explaining these phenomena of spatial diffusion or contagion, the authors followed the example examined by Elkins *et al.* (2006) who argued that the spread of Bilateral Investment Treaties (BTI) is driven by international competition among potential host countries for Foreign Direct Investment (FDI). In order to assess the strength of competition and cultural emulation, Elkins *et al.* (2006) constructed, on the basis of the theoretical framework traced in a previous work (Elkins and Simmons, 2005), a series of spatial lags able to capture the behaviour of neighbours.

The spatial lag is generally given by a $N \times N \times T$ spatial weight matrix W that maps the distances between units for each year, multiplied for the lagged dependent variable for all countries other than a generic country i.

In a dyadic setting, as argued also by LeSage and Pace (2008), the unit is represented by the country-pair ij so that the weight is associated to each couple of observations. In particular, the probability that people move from country i to country j depends on the weighted sum of all other migration flows existing between source countries k and destination countries m, where the unique restriction is that the excluded dyad is exactly the specific couple ij. This means that the dyad which contributes to the contagion effect may contain alternately the same source country i or the same destination country j. Using the formalization adopted in Neumayer and Plümper (2010) and ignoring the time dimension, the spatial variable is:

$$y_{ij} = \rho \sum_{km \neq ij} \omega y_{km} + \varepsilon_{ij}$$

where ρ is the spatial autoregressive parameter associated to the spatial lag $\sum_{km \neq ij} \omega y_{km}$ which consists of a spatial weighting matrix ω and the temporally lagged value of the dependent variable in all country-pairs km, y_{km} .

A part from the traditional spatial techniques here considered, Kapetanios *et al.* (2014, hereafter KMS) developed a general econometric modeling framework which allows to take into account the influence of peer effects. Based on the idea that economic agents consider the views and behaviours of those around them and aggregate them in order to form their own view, this approach allows to introduce in the model an interaction term of neighbouring units. The key feature in this class of models is that the computation of this spatial term is unit-specific; it takes unit-specific aggregate which determines the closeness of units through a threshold mechanism. The purpose of this approach is twofold: it allows cross-sectional dependence to arise endogenously and it captures different kind of links within units. In this context, the methodology proposed by KMS (2014) represents an innovation in the empirical literature of migration and so it constitutes an alternative to the usual spatial econometric techniques in capturing the diffusion or contagion effect which characterizes the connection across countries.

To formalize this idea, we consider a sample of N units given by country-pairs ij, observed for T periods and we specify the spatial variable as:

$$y_{ijt} = \frac{1}{h_{ijt}} \sum_{km \neq ij} \ell(|d| \leq r) y_{km(t-1)}$$

with $h_{ijt} = \sum_{km \neq ij} \ell(|d| \leq r)$, where y_{ijt} represents the migration flow within a specific country-pair, denoted with ij, at time t which is influenced in some nonlinear fashion by the cross-sectional average of lagged flows of a selection of neighbouring country-pairs, $y_{km(t-1)}$, identified through a specific value assigned to the threshold parameter r according to a constraint applied on a certain measure of distance d.

The concept of neighbourhood is associated to distance in geographic terms, since we suppose that similar behaviour in migration flows are driven by similar distances between different country-pairs.

Once considered the problem of multilateral resistance and contagion effects, it's important to evaluate how to estimate the gravity equation. The primary and more popular estimation methodology is the ordinary least square (OLS) estimator.

The general gravity model for migration is given by:

$$F_{ijt} = e^{\zeta_{it}} e^{\chi_{jt}} X_{ijt}^{\beta_1} D_{ij}^{\beta_2} \tilde{F}_{ijt}^{\beta_3} \varepsilon_{ijt} \tag{1}$$

where ζ_{it} and χ_{jt} are destination and origin specific fixed effects, $X_{ijt}$is a vector containing the covariates of interest, such as the "mass" of source and destination countries, D_{ij} is a distance variable, $\tilde{F}_{ijt}$ is the spatial interaction term and ε_{ijt} is an error factor assumed to be statistically independent of the regressors.

The traditional approach of estimation is to consider the log-linearized specification of equation (1) and to estimate the parameters of interest by ordinary least squares (OLS):

$$\ln(F_{ijt}) = \zeta_{it} + \chi_{jt} + \beta_1 \ln X_{ijt} + \beta_2 \ln D_{ij} + \beta_3 \ln \tilde{F}_{ijt} + \ln \varepsilon_{ijt} \tag{2}$$

as underlined by Silva and Tenreyro (2006), Tenreyro (2007) and Egger and Staub (2016), the OLS estimation will only be consistent if and only if the following condition is fulfilled:

$$E[\ln \varepsilon_{ijt} | \zeta_{it}, \chi_{jt}, \ln X_{ijt}, \ln D_{ij}, \ln \tilde{F}_{ijt}] = 0$$

It is very unlikely that ε_{ijt} is independent of the countries' mass and of the bilateral distances, because according to the economic theory the error term is generally heteroskedastic. Therefore, the OLS method could lead to biased estimates.

Secondly, in the case of country-pairs with zero migration flows, these are dropped out of the sample as a result of logarithmic transformation and this could be a source of additional biases.

Therefore, the Poisson Pseudo-Maximum Likelihood (PPML) represents a valid alternative to OLS as suggested by Silva and Tenreyro (2006) for consistent estimation of the parameters of the gravity model and additionally, it solves the problem of zero migration flows because the dependent variable is taken in levels, instead of logarithmic terms.

In particular, the parameters in equation(1) can be estimated using the fact that the conditional expectation of F_{ijt}, with respect to the set of covariates $Z_{ijt} = (\zeta_{it}, \chi_{jt}, X_{ijt}, D_{ij}, \tilde{F}_{ijt})$, can be written as the following exponential function:

$$E(F_{ijt}|Z_{ijt}) = \exp[\zeta_{it} + \chi_{jt} + \beta_1^{'} \ln X_{ijt} + \beta_2^{'} \ln D_{ij} + \beta_3 \ln \tilde{F}_{ijt}] \tag{3}$$

This implies the following:

$$F_{ijt} = \exp[\zeta_{it} + \chi_{jt} + \beta_1^{'} \ln X_{ijt} + \beta_2^{'} \ln D_{ij} + \beta_3^{'} \ln \tilde{F}_{ijt}] + \varepsilon_{ijt} \tag{4}$$

where ε_{ijt} is the remainder error term in additive form.

This approach has been supported also by other authors. For instance, Henderson and Millimet (2008) tried to estimate gravity models in both log-linearized terms and in levels, comparing parametric and non-parametric structures. They show that, despite the added flexibility of nonparametric models, the PPML specification in its parametric form offers equally or more reliable in-sample and out-of-sample forecasts.

In this work, we propose an application of the methodologies described above on the RUM model specified by Beine *et al.* (2016) and discussed also by Ramos (2017), applied to a dataset which extends the one used by Ramos and Suriñach (2017) in terms of time. In their paper, they analyzed past and future trends of migration flows between the European Union (EU) and the European Neighbouring countries (ENC), considering the fundamental determinants which attract extra-EU citizens towards EU members. In this paper, we consider the determinants of unidirectional migration flows from ENC to EU countries over a period of 18 years.

Therefore, section 2 focuses on the RUM model specification for gravity models of migration, the PPML estimation method, considering the drawbacks in the use of the traditional least square estimator, and the KMS contribution in determining the diffusion effect. Section 3 presents the ENC-EU dataset and provides a discussion of the main empirical results, while section 4 offers some concluding remarks.

2. The Model

2.1 The RUM Model

To the best of our knowledge, the Random Utility Maximization (RUM) model represents the more considerable attempt to draw the micro-theoretical basis for the estimation of gravity equations in the context of migration analysis.

Following Beine *et al.* (2016) and Ramos (2017), the RUM model of migration describes the utility that a generic agent h located in country i at time $t-1$ gets by moving towards country j belonging at a specific choice set B at time t as:

$$U_{hijt} = w_{ijt} - c_{ijt} + \varepsilon_{hijt} \tag{5}$$

where w_{ijt} represents a deterministic component of utility, c_{ijt} is the cost of moving from i to j between time $t-1$ and t and ε_{ijt} is the individual-specific stochastic error term.

The choice about the distribution of the stochastic error term ε_{ijt} affects the expected probability of individual s to maximize her utility in moving from her origin country to another one. Specifically, if we assume that the error term is i.i.d. as in Grogger and Hanson (2011), and we consider the axiom of irrelevance of alternative destination effects, as shown in McFadden (1974), we rewrite the expected probability to migrate from country i to country j as:[1]

$$E(p_{ijt}) = \frac{e^{w_{ijt}-c_{ijt}}}{\sum_m e^{w_{jmt}-c_{jmt}}} \tag{6}$$

where m represents any destination country belonging to the choice set B.

Since the gross migration flow F_{ijt} from country i to country j at time t is:

$$F_{ijt} = p_{ijt} s_{it}$$

where s_{ijt} is the stock of population residing in the source country i at time t, we can derive from equation (6) the expected gross migration flow as the product of the stock of population at time t by the

[1] McFadden (1974) consider an axiom of independence of irrelevant alternatives introduced by Luce (1959) "which states that the relative odds of one alternative being chosen over a second should be independent of the presence or absence of unchosen third alternatives". He shows that in empirical analysis, when the utility function is additively separable as in our case, the multiple choice of selection probabilities can be written in terms of binary odds as represented in equation (30).

expected probability $E(p_{ijt})$:

$$E(F_{ijt}) = \frac{e^{w_{ijt}-c_{ijt}}}{\sum_m e^{w_{jmt}-c_{jmt}}} s_{it} \tag{7}$$

Assuming that the deterministic component of utility, w_{ijt} does not vary with the origin country i, equation (7) can be reduced to resamble the traditional gravity equation as:

$$E(F_{ijt}) = \phi_{ijt} \frac{x_{jt}}{\Omega_{it}} s_{it} \tag{8}$$

where: $x_{jt} = e^{w_{jt}}$ represents the attractiveness of destination country j at rime t, $\phi_{ijt} = e^{-c_{ijt}}$ is the accessibility of destination country k for migrants coming from i at time t and it has a negative impact on migration flows. The term $\Omega_{it} = \sum_m e^{w_{jt}-c_{ijt}} x_{mt}$ represents, according to Small and Rosen (1981), the exponential vale of the expected utility for agent i of choosing a different destination location or not migrating, that is opting for the source country i. Therefore, Ω is strictly related to the concept of multilateral resistance to migration, since migration flows could be influenced not only by the attractiveness and the cost to move to the destination country j but also by the attractiveness of alternative target countries.

The stochastic formulation of the above gravity model in (8) implies an error term ε_{ijt}, with $E(\varepsilon_{ijt}) = 1$, so that:

$$F_{ijt} = \phi_{ijt} \frac{x_{jt}}{\Omega_{it}} s_{it} \varepsilon_{ijt} \tag{9}$$

Looking at the estimation issue, as highlighted in Beine *et al.* (2016), the gravity model derived from the RUM model can be estimated in two different ways. In a first case, the dependent variable is the level of bilateral gross migration flows F_{ijt}. In a second case we can look at a ratio q_{ijt} of choice probabilities defined as the ratio at time t between the expected migrants from country i to country j, $E(F_{ijt})$, and the expected stayers of country i, $E(F_{iit})$, that is:

$$q_{ijt} = \frac{E(F_{ijt})}{E(F_{iit})} = \phi_{ijt} \frac{x_{jt}}{x_{it}}$$

where x_{jt} and x_{it} represent, respectively the pull factors of destination country j and source country i.

Using q_{ijt} as dependent variable means estimating the following OLS model where the regressors are expressed in logs:

$$q_{ijt} = \ln\phi + \ln x_{jt} - \ln x_{it} + \ln\left(\frac{\varepsilon_{ijt}}{\varepsilon_{iit}}\right) \tag{10}$$

so that, the ratio of choice probabilities, q_{ijt}, is positively related to the attractiveness and accessibility of country j but negatively related to the convenience not to migrate. The focus of this specification should relate the error term, since the OLS estimator is consistent only if $E[\ln(\varepsilon_{ijt}/\varepsilon_{iit})]$=0

In the original RUM model, we assume that the error term ε_{ijt} is i.i.d. and that its expectation $E(\varepsilon_{ijt}) = 1$, but this doesn't imply that the OLS estimator of the log-linearized model is not biased. As explained by Santos Silva and Tenreyro (2006), the expectation on the logarithm of a random variable depends both on its mean and on higher-order moments of the distribution. In particular, if ε_{ijt} is log-normally distributed, with $E(\varepsilon_{ijt}) = 1$ and variance σ^2_{ijt}, it follows that $E[\ln\varepsilon_{ijt}] = -\frac{1}{2}\ln(1+\sigma^2_{ijt})$. Therefore, if we consider that the error term is heteroskedastic, since its variance depends on the regressors of the gravity equation, this implies that the expectation on $\ln\left(\varepsilon_{ijt}/\varepsilon_{iit}\right)$ also depends on the regressor. Hence, the traditional OLS method leads to inconsistent and biased estimates. Another problem occurs when migration flows between different pairs are exactly equal to zero. In the presence of zero-observations of the dependent variable, when we take the logarithm of migration flow, in order to apply the OLS estimator, we should consider a reduced sample of non-zero observations, which could be another source of misleading results. In particular, the exclusion of zero-observations could lead to the so called sample selection bias that results from a non-random sampling.

The proposed solution is to estimate the model with Poisson Pseudo-Maximum Likelihood (PPML), a technique which requires weaker assumptions than OLS in order to get consistent estimates. In general, the Poisson distribution is used for count data models, but it could be also applied more generally to nonlinear model such as gravity, since we are dealing with a pseudo-maximum likelihood (PML) estimator. In the PML estimator, "we maximize the log-likelihood function by equating its derivatives to zero" (Greene, 2003). As explained in Santos Silva and Tenreyro (2006), the consistency of the PPML results depends solely on the correct specification of the gravity equation: *i.e.* on the regressors included. In particular, the PPML estimator can accommodate for the presence of fixed effects as in the traditional OLS regression, considering origin-time dummies and destination-time dummies in order to account for MLR. By construction, the dependent variable is expressed in levels, thus this estimator is able to deal with the presence of many zero values of bilateral gross migration flows. The PPML estimator performs well even in presence of zeros in the data, and allows to overcome a strong limitation of the traditional OLS. Indeed, in OLS, since the logarithm of zero is not defined, all those observations equal to zero are dropped from the dataset. Severe biases could arise from an analysis of the reduced dataset, so the use of PPML is highly recommended in this case.

Apart from being consistent with heterogeneity in the propensity to migrate and in the presence of zero values in migration flows, the PPML displays another desirable property: the interpretation of the coefficients is undemanding and follows the classical interpretation of OLS coefficients.

In the next section, we will explain the origin of PPML estimator and its basic assumptions, which make it a desirable approach for dealing with gravity models.

2.2 The Poisson Pseudo-Maximum Likelihood (PPML) Specification

As explained by Santos Silva and Tenreyro (2006), when we consider constant-elasticity models as the economic theory of gravity models suggests, the link between the generic response y_i and the covariate x_i is of the form:[2]

$$y_i = \exp(x_i\beta) \tag{11}$$

where the function $\exp(x_i\beta)$ is interpreted as the conditional expectation of y_i given x_i. Since this relation $y_i = \exp(x_i\beta)$ holds on average but not for each observed i, an error term ε_i given by $y_i - \exp(x_i\beta)$, is associated with each i. Therefore, the stochastic model can be formulated in two equivalent representations with additive or multiplicative error terms, as:

$$y_i = \exp(x_i\beta) + \varepsilon_i = \exp(x_i\beta)\,\eta_{ij} \tag{12}$$

with $y_i \geq 0$ and the error terms which are mean-independent of the covariate, *i.e.* respectively: $E(\varepsilon_i|x_i) = 0$ and $E(\eta_i|x_i)$=1, where $\eta_i = 1 + \varepsilon_i/\exp(x_i\beta)$.

If we consider the OLS estimate of the model with multiplicative errors by taking natural logarithms of both sides of equation (12) :

$$\ln y_i = x_i\beta + \ln \eta_i$$

[2]For reasons of simplification, we consider a single covariate x_i

in order to obtain consistent estimates of β, it is necessary that the conditional expectation of the error in logarithmic terms does not depend on the covariate, *i.e.* $E(\ln \eta_i|x_i) = 0$. This condition is met only if $\varepsilon_i = \exp(x_i\beta)v_i$ where v_i is a random variable statistically independent of x_i, alongside $\eta_i = 1 + v_i$. As stated in Santos Silva and Tenreyro (2006), when the condition of statistically independence of η_i is matched, the conditional variance of the response variable, $V[y_i|x_i]$ is proportional to $\exp(2x_i\beta)$. However, it's not possible to derive this information from the economic theory, indeed we can infer some information from the data. Since we said above that y_i can assume non-negative values, it means that if the conditional expectation on the response y_i tends to zero, the associated probability also approaches 0 and the conditional variance $V[y_i|x_i]$ tends to vanish. On the contrary, if the conditional expectation is large, the correspondent variance is high. The heteroskedasticity of the conditional distribution of the error term can be a nontrivial problem in the estimation of this type of models. Therefore, the log linear model estimated trough OLS could lead to seriously misleading results, due to the unrealistic assumption that the conditional expectation of $\ln y_i$ is a linear function of the covariates.

A possible solution to this is the use of a non-linear least squares estimator (NLS) as done by Frankel and Wei (1993) in the case of a constant-elasticity model for international trade. In particular the NLS estimator is defined as:

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} [y_i - \exp(x_i\beta)]^2$$

and the set of first-order conditions (FOC) is given by: $\sum_{i=1}^{n} \left[y_i - \exp(x_i\hat{\beta}) \right] \exp(x_i\hat{\beta}) x_i = 0$.

The NLS estimator presents a drawback: it ignores heteroskedasticity. This is due to the fact that the set of FOC gives more weight to noisier observations, that is to the observations where $\exp(x_i\hat{\beta})$ is large, hence a small number of observations of this type could heavily affect the efficiency of the NLS estimator. The problem could be solved by using a weighted non-linear least square which will be possible if and only if the form of the conditional variance $V[y_i|x_i]$ is known. Robinson (1987) proposed a nonparametric generalized least squares estimator which account for heteroskedasticity, which has been applied by O'Hara and Parmeter (2013) for handling heteroskedasticity to data on professor rankings obtained from *RateMyProfessors.com*. However this approach has never been adopted as a standard tool by researchers.

Therefore, Santos Silva and Tenreyro (2006) discuss in their paper about a more efficient estimator than the NLS without using nonparametric regression and assuming that the conditional variance $V[y_i|x_i]$ is simply proportional to the conditional mean $E(y_i|x_i)$. Under this hypothesis and as explained further, the considered set of FOC is:

$$\sum_{i=1}^{n} \left[y_i - \exp(x_i\tilde{\beta}) \right] x_i = 0 \tag{13}$$

The estimator based on this set of equation is called Poisson Pseudo-Maximum Likelihood (PPML) estimator, it gives the same weight to all observations under the assumption of proportionality between $E(y_i|x_i)$ and $V[y_i|x_i]$. It is more efficient than the standard NLS estimator and moreover it needs only a correct specification of the conditional mean $E(y_i|x_i) = exp(x_i\beta)$ in order to be consistent.

The PPML regression is considered as a special case of nonlinear models in the class of Generalized Linear Models (GLM). In order to understand how it works, we propose below a description of the structure of these models as proposed in the book of Farhmeir and Tutz (2013) and in the paper of Egger and Staub (2016).

The Generalized Linear Models (GLM) are based on the fact that:

i) the covariates x could be linked in a nonlinear fashion with the response variable y;

ii) the variance of the error term could be not constant (heteroskedastic);

iii) the distribution of the response function could be not normal.

Specifically, GLM specification considers two assumptions:

1) a *distributional assumption*: we assume that y_i are conditionally independent, given the covariate x_i, and the conditional distribution of y_i belongs to a Linear Exponential Family (LEF), which is a family of probability measures, with conditional expectation $E(y_i|x_i) = \mu_i$;

2) a *structural assumption*: the link between the linear predictor $\gamma = x_i\beta$ and the expectation μ is not linear. It is expressed by an invertible function g called link function, able to connect γ and μ in the following way:

$$\mu = g^{-1}(x_i\beta) = g^{-1}(\gamma)$$

where $\gamma = g(\mu)$, e.g. for the Poisson $\gamma = log(\mu)$, so that $\mu = \exp(x_i\beta)$.

Therefore, a GLM is characterized by the type of the exponential family (EF), the response or link function g and the covariate x.

The generic density function related to these LEF is:

$$f(y_i|\theta_i, \psi, \omega_i) = \exp\left\{ \left(\frac{\omega_i}{\psi}(y_i\theta_i - b(\theta_i) \right) + c(y, \psi, \omega_i) \right\}$$

where θ is the canonical or natural parameter, ψ is the dispersion parameter, $b(\cdot)$ and $c(\cdot)$ are specific function depending on the type of EF and ω is a weight that for ungrouped data, where $i = 1, \ldots, n$, is equal to 1.

For such a random variable, assuming $\omega_i = 1$, the defined log-likelihood will be:

$$l(\theta_i) = \frac{y_i\theta_i - b(\theta_i)}{\psi} + c(y_i, \psi)$$

and the score function, defined as the partial derivative of the log-likelihood function with respect to the set of canonical parameters θ_i, is:

$$s(\theta_i) = \frac{\partial l(\theta_i)}{\partial(\theta_i)} = \frac{y_i - b^{'}(\theta_i)}{\psi}$$

Taking the expected score and imposing it equal to zero, we obtain the following FOC:

$$E\left[s(\theta_i)\right] = \frac{E(y_i) - b^{'}(\theta_i)}{\psi} = 0$$

The set of parameters θ_i is a function of the mean μ_i , *i.e.* $\theta_i = \theta(\mu_i)$, which, according to the chosen EF is determined through the relation $\mu_i = b'(\theta_i)$. Furthermore, as the structural assumption on GLM suggests, the mean of the response y_i, μ_i, is specified as a function of a linear index of the considered set of covariates $x_i\beta$, that is $\mu_i = g^{-1}(x_i\beta)$. The form of the link function for the Poisson distribution applied on gravity is a log-link, as previously defined.

The second important result of the likelihood theory, as suggested in Egger and Staub (2016), is the derivation of the variance function. By defining the Hessian of a LEF random variable as follows:

$$H(\theta_i) = \frac{\partial^2 l(\theta_i)}{\partial \theta_i^2} = \frac{-b^{''}(\theta_i)}{\psi}$$

and considering the information matrix equality $V\left[s(\theta)\right] = -E\left[H(\theta)\right]$, we can write:

$$\frac{E\left[y_i - b^{'}(\theta_i)\right]^2}{\psi^2} = \frac{b^{''}(\theta_i)}{\psi}$$

which is equivalent to $V(y_i) = b^{''}(\theta_i)\psi$.

The set of parameters θ is estimated by maximizing the sample pseudo-likelihood function:

$$l_{sample}(\beta) = \sum_{i=1}^{n} \frac{y_i\theta_i - b(\theta_i)}{\psi} + c(y_i, \psi)$$

with corresponding score:

$$s_{sample}(\beta) = \sum_{i=1}^{n} \frac{y_i - b^{'}(\theta_i)}{\psi} \frac{\partial \theta_i}{\partial \beta} = \frac{y_i - b^{'}(\theta_i)}{\psi} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta} = \frac{y_i - b^{'}(\theta_i)}{b^{''}(\theta_i)\psi} \frac{\partial \mu_i}{\partial \beta}$$

since $\mu_i = b^{'}(\theta_i)$. Thus, the FOC for the above score are:

$$\sum_{i=1}^{n} \frac{[y_i - E(y_i)]}{V(y_i)} \frac{\partial E(y_i)}{\partial \beta} = 0$$

Therefore, looking at the Poisson case and following the pioneer work of Gourieroux *et al.* (1984), the objective (score) function corresponding to the Poisson family is given by:

$$-\sum_{i=1}^{n} \exp(x_i\beta) + \sum_{i=1}^{n} y_i x_i \beta$$

and the relative pseudo-likelihood equation (FOC) which leads to the correct estimation of β is:

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

Given these fundamental results of GLM theory, we can apply them to the context of migration gravity models.

Let us define the following migration gravity equation as in equation (1):

$$F_{ijt} = e^{\zeta_{it}} e^{\chi_{jt}} X_{ijt}^{\beta_1} D_{ij}^{\beta_2} \tilde{F}_{ijt}^{\beta_3} \varepsilon_{ijt} \tag{14}$$

where F_{ijt} represents the response variable relative to migration flows from a generic source country i to a destination country j; ζ_{it} and $\chi_{jt}$are respectively source-time fixed effects and destination-time fixed effects, considered in order to account for MLR; X_{ijt} is a vector containing the covariates of interest, such as the "mass" of source and destination countries; D_{ij} identifies the cost of moving, *i.e.* geographical distance; $\tilde{F}_{ijt}$ is the spatial covariate which captures some interaction effects between country-pairs and ε_{ijt} is an error factor assumed to be statistically independent with respect the regressors.

We consider the conditional expectation of migration flows F_{ijt} with respect to the set of regressors $Z_{ijt} = (\zeta_{it}, \chi_{jt}, X_{ijt}, D_{ij}, \tilde{F}_{ijt})$:

$$E(F_{ijt}|Z_{ijt}) = \exp[\zeta_{it} + \chi_{jt} + \beta_1' \ln X_{ijt} + \beta_2' \ln D_{ij} + \beta_3 \ln \tilde{F}_{ijt}] \tag{15}$$

and the following stochastic model which has to be estimated through the PPML estimator:

$$F_{ijt} = \exp[\zeta_{it} + \chi_{jt} + \beta_1' \ln X_{ijt} + \beta_2' \ln D_{ij} + \beta_3' \ln \tilde{F}_{ijt}] + \varepsilon_{ijt}$$

$$= \exp[\zeta_{it} + \chi_{jt} + \beta_1' \ln X_{ijt} + \beta_2' \ln D_{ij} + \beta_3' \ln \tilde{F}_{ijt}]\eta_{ijt} \tag{16}$$

Once the gravity specification has been described, a special focus has to be given to the spatial component $\tilde{F}_{ijt}$. It represents an added covariate, able to capture a diffusion effect in migration policies, modeled through an alternative spatial technique suggested by KMS (2014). The next two sections will be devoted to the examination of the spatial effects involved in gravity.

2.3 The Diffusion Effect

The introduction of spatial effects in the gravity equation has been achieving an increasing attention in explaining the effects of migration networks. In particular, in order to capture the multilateral resistance frictions, the identification of endogenous peer effects could provide a useful prediction of the influence of networks on migration flows. As defined in the introduction, multilateral resistance to migration is related to the influence of third countries in determining migration flows within a specific country-pair. The use of specific fixed effects, as demonstrated in the literature on gravity, attempts to control for MLR. In particular, Ortega and Peri (2013) introduced in their RUM model (estimated with OLS) origin-fixed effects which capture all terms that vary across origin and year but not across destinations. Beine and Pearsons (2015) proposed otherwise the use of destination-year fixed effects, since migration policies are defined at target level. Moreover, apart from the influence of possible destination alternatives, Neumayer and Plümper (2010a) focused their attention on the possible spatial interdependence between different pairs of countries, also defined as dyads, due to policy choices. They developed a unified framework where all possible forms of spatial dependence could arise with the direct inclusion of spatial effects in dyadic data. This avoids employing Bayesian hierarchical or random effects models as suggested, conversely, by King (2001) as starting methodological point in the study of international relations.

Specifically, Neumayer and Plümper (2010a) distinguish different types of contagion or diffusion effects in a directed dyadic data setting. The dyad is a specific unit of analysis formed by a pair of two political units. It could be undirected when it is not possible to operate a distinction between the two members of the pair, or directed, if instead it is well known that, in the couple ij, the interaction initiates with i and is directed towards j. Thus, in the last case, the distinction between source/origin and destination/target countries of the dyad is clear. If we analyze directed dyadic data, there are five options of modeling spatial interdependence.

A first candidate is the directed dyad contagion effect which describes a situation where migration flows within the pair ij could be affected by the weighted sum of all other migration flows between source countries k and destination countries m, where alternately i may coincide with k or the target j could be equal to m. This situation is presented with the following notation:

$$F_{ijt} = \rho \sum_{km \neq ij} wF_{kmt} + \varepsilon_{ijt}$$

where ρ is the spatial lag parameter to be estimated and ω is a time-invariant spatial weight matrix (Ramos, 2017). In this sense, "the aggregate policy choices of other, similar dyads matters" (Neumayer and Plümper, 2010a).

The residual four forms of contagion relate a subset of observed countries.

When flows between the country-pair ij depend only on the aggregate flows of source countries different from i, an aggregate source contagion effect is defined:

$$F_{ijt} = \rho \sum_{k \neq i} \sum_{m} wF_{kmt} + \varepsilon_{ijt}$$

if the flow depends on the aggregate flows of destination countries, despite of sources, we define an aggregate target contagion as:

$$F_{ijt} = \rho \sum_{m \neq j} \sum_{k} wF_{kmt} + \varepsilon_{ijt}$$

Moreover, if we consider that the spatial interdependence is due to the choice of sources in relation to a specific dyad, we define a specific source contagion term:

$$F_{ijt} = \rho \sum_{k \neq i} wF_{kmt} + \varepsilon_{ijt}$$

while when other targets m influence the migration flows between j and i , we have a specific target contagion effect defined as:

$$F_{ijt} = \rho \sum_{m \neq j} wF_{kmt} + \varepsilon_{ijt}$$

The choice on the right specification of the spatial interaction term belongs to the researchers according to the theory they want to test.

A trivial point in the definition of the spatial lagged variable is also given by the specification of the weight matrix. As underlined in Plümper and Neumayer (2010), the definition of the functional form of the weighting matrix when there isn't a clear dichotomous measure of connectivity such as geographical distance could give different results. At the same time, the common practice of row-standardization could alter the estimating results. Thus, the theoretical framework should guide the researcher to adopt a correct functional form of the weighting matrix.

What we propose hereafter is a method which defines the spatial variable overcoming the problem of row-standardization. Based on the assumption that "near things are more related than distant things" (Tobler,1970), it's possible to define the weights by considering an indicator function which equals 1 when the absolute value of a specified difference lies below a certain threshold. Hence, the spatial term is identified by following the KMS (2014) approach considering a directed dyad contagion effect, related to the influence exerted on a generic country-pair ij by the aggregate of other similar dyads.

2.4 Spatial Term Definition

Based on the hypothesis that country choices are often not independent of actions implemented by other countries, the inclusion of a variable able to capture the spatial interdependence in the gravity model could give relevant insights. In particular, the conclusion of such agreements between countries often generates externalities, which could be positive or negative depending on the produced effects. In our application, where we consider the migration flow that originates from EU's neighbours and is directed towards EU members, the remarkable phenomenon which could lead to a positive externality, is the relation that occurs between a given country-pair and all the other dyads in the set of considered countries. Specifically, we expect that similar dyads act in the same way producing an emulation or globalization effect. The similarity between different dyads could be easily captured by using a threshold mechanism associated to a kind of distance measure between dyads.

Motivated by this consideration, we construct a spatial variable inspired to the relevant work of KMS (2014). Originally, they developed a class of nonlinear panel data models that allows to control the problem of cross-sectional dependence (CSD). Differently, in our case study, we borrow the approach suggested in KMS to account for CSD in order to model the interdependence among country-pairs in the pattern of migration flows.

According to KMS (2014) approach, we can define a function which endogenously determines the neighbourhood or similarity between units according to a specific concept of distance. Since we are interested in considering a directed dyad contagion effect, given by the aggregate behaviour of "neigbouring" or similar dyads, we can define a spatial lagged variable by using a threshold regime selection mechanism.

Let $F_{km(t-1)}$ be the lagged migration flow associated to the generic country pair km, where k denotes the source country and m the corresponding destination country, $d_{ij} = \ln D_{ij}$ be the logarithm of the geographical distance between country i and the target j, similarly for d_{km} for the pair km, we define the spatial variable $\tilde{F}_{ijt}$ as:

$$\tilde{F}_{ijt} = \frac{1}{h_{ijt}} \sum_{km \neq ij} \ell \left(|d_{ij} - d_{km}| \leq r \right) F_{km(t-1)} \tag{17}$$

with $h_{ijt} = \sum_{km \neq ij} \ell \left(|d_{ij} - d_{km}| \leq r \right)$, cross-sectional average of $F_{km(t-1)}$of those country-pairs km similar in terms of distances between the source and destination countries to the country-pair ij, according to a specific threshold r. The threshold parameter r is endogenously determined by means of a grid search algorithm and for each value of r different estimation of $\tilde{F}_{ijt}$ are computed. The symbol ℓ introduces an indicator function which is equal to 1 if the absolute difference of the distances associated with the two country pairs ij and km lies under a given threshold r. Thus, the relevant lagged migration flows are the one for which the condition described through the indicator function is matched. In particular, we evaluate that a similar behaviour in migration policies is associated to those country-pairs characterized by a similar distance within the countries forming the pair. In other words, considering the two generic country-pairs ij and km, if the distance between i and j is more or less equal to the distance between k and m and, more specifically, if the absolute difference between the two distances is below the value of the r parameter, the migration flow at time $t-1$ of unit km has a relevant influence on average on the migration flow for the country-pair ij. This idea represents the aggregate diffusion effect on migration flows between i and j given by the average of flows in similar country-pairs km.

The value of r is obtained by least squares cross-validation, that is by minimizing the sum of squared residuals given hereunder:

$$V(\rho, r) = \min_{\rho,r} \frac{1}{N}\frac{1}{T}\sum_{ij}\sum_{t}\hat{\varepsilon}_{ijt}(\rho, r)$$

$$= \min_{\rho,r}\left[F_{ijt} - \frac{\rho}{h_{ijt}}\sum_{km\neq ij}\ell\left(|d_{ij} - d_{km}| \leq r\right) F_{km(t-1)}\right]^2$$

The minimization process allows to obtain the correct estimate of the spatial variable $\tilde{F}_{ijt}$ which enters the PPML estimator in logarithm by construction.

Therefore, the proposed gravity model is:

$$F_{ijt} = \exp[\zeta_{it} + \chi_{jt} + \beta_1^{'} \ln X_{ijt} + \beta_2^{'} \ln D_{ij} + \beta_3^{'} \ln \tilde{F}_{ijt}] + \varepsilon_{ijt}$$

$$= \exp[\zeta_{it} + \chi_{jt} + \beta_1^{'} \ln X_{ijt} + \beta_2^{'} \ln D_{ij} + \beta_3^{'} \ln \tilde{F}_{ijt}]\eta_{ijt} \quad (18)$$

where

$$\tilde{F}_{ijt} = \frac{1}{h_{ijt}} \sum_{km \neq ij} \ell\left(|d_{ij} - d_{km}| \leq r\right) F_{km(t-1)} \tag{19}$$

and considering $x_{ijt} = \ln X_{ijt}$, $d_{ij} = \ln D_{ij}$, $\tilde{f}_{ijt} = \ln \tilde{F}_{ijt}$ it is rewritten as:

$$F_{ijt} = \exp[\zeta_{it} + \chi_{jt} + \beta_1' x_{ijt} + \beta_2' d_{ij} + \beta_3 \tilde{f}_{ijt}] + \varepsilon_{ijt}$$

$$= \exp[\zeta_{it} + \chi_{jt} + \beta_1' x_{ijt} + \beta_2' d_{ij} + \beta_3 \tilde{f}_{ijt}] \eta_{ijt} \tag{20}$$

The spatial variable computed through the KMS approach presents a value added with respect to the traditional spatial techniques: the diffusion effect is the outcome of the influence caused only by those country-pairs similar between them, where similarity is expressed as a function of distance.

3. Empirical Results

3.1 Data Description

For this empirical study, data have been collected from different data sources listed below:

1) the OECD International Migration Database provides tables with annual series on migration flows and stocks of foreign-born and foreigners of OECD countries;

2) the World Bank provides a lot of global development data including population stock and GDP values of countries all over the world;

3) the CEPII database provides several data about gravity models, specifically it makes available data on geographical variables, such as bilateral distances (in kilometers) measured using the geographical coordinates of countries' capital cities, community of border and languages, and colonial history. The CEPII also includes in the gravity dataset measures of GDP and population, however we decided to take into consideration the estimates provided by the World Bank Database since it contains longer time series on the variables of interest.

It is worth noting that the object of the following empirical study has been the determinant of migration flow in the dyadic framework where the source countries are identified with the so called European Neighbouring Countries (ENC), while the destination countries are European nations. The ENC are a group of 16 countries which have established a cultural cooperation with the EU, which goes under the name of European Neighbourhood Policy (ENP). This partnership governs the EU's relation with the following Eastern and Southern countries: Algeria, Armenia, Azerbaijan, Belarus, Egypt, Georgia, Israel, Jordan, Lebanon, Libya, Moldova, Morocco, Palestine, Syria, Tunisia and Ukraine. These countries cooperate with EU members in order to foster stabilization, security and prosperity in line with a global strategy. In addition to the aforementioned countries, recently, within the Euro-Mediterranean Partnership or Union for the Mediterranean (UfM), some other countries are playing an important role, such as: Albania, Bosnia and Herzegovina, Mauritania, Monaco, Montenegro and Turkey. Due to the lack of data about migration flows, in our dataset, within the group of source countries, we do not consider Azerbaijan, Bosnia and Herzegovina, Mauritania, Monaco, Montenegro and Palestine. Within the EU members we took into consideration a set of 11 countries with available data along the chosen period, *i.e.* Austria, Denmark, Finland, France, Germany, Italy, Luxembourg, Netherlands, Norway, Spain and Sweden.

Based on the closeness relation between the EU and the ENCs and on the availability of data, we propose an analysis of the unidirectional migration flows from ENC to EU, considering a panel dataset for the years 2000-2017.

Looking at inflows from ENCs, the favourite destination countries seem to be Austria, France, Germany, Italy, Netherlands and Spain. In Table 6, we report the contribution of the main source countries for the mentioned EU members.

Main destination Countries (EU members)	**Main source countries (ENCs)**
Austria	Syria (43,13%); Turkey (21,57%)
France	Algeria (31,18%); Morocco (26,85%); Tunisia (16,68%)
Germany	Syria (40,90%); Turkey (18,02%); Albania (7,98%); Ukraine (7,02%)
Italy	Morocco (27,08%); Albania (26,56%); Ukraine (13,57%); Egypt (13,31%)
Netherlands	Syria (61,31%); Morocco (5,88%)
Spain	Morocco (64,38%); Ukraine (12,55%); Algeria (7,73%)

Table 1: Percentage of total migrants in 2017 above 5% on ENC total inflows

* In Norway and Sweden, respectively, the 81,39% and the 76,03% of total ENC migrants comes from Syria

Source: Own elaboration from OECD International Migration Database

As stated above, we considered as dependent variable the annual series of migration flows in levels from the OECD International Migration Database, where immigrants are identified using the foreign-born criteria, and as basic regressors the population stocks and the GDP of both source and destination countries in logarithmic terms. Specifically, we take the GDP per capita value (constant 2010 US dollars), accordingly to the literature on gravity models.

Therefore, our potential dataset includes 3168 observations given by 176 country-pairs (16 source and 11 destination countries) over a period of 18 years.

3.2 Main Results

There are several specifications of gravity models for migration, according to different determinants of flows. As seen in the introduction, based on Newton's gravity law, the economic theory developed constant-elasticity models, with the particular case of RUM models. As Santos-Silva and Tenreyro (2006) pointed out, the most common practice to estimate gravity with log-linear models could lead to misleading results for the inconsistency of the OLS estimator in the presence of heteroskedasticity and of zero values in migration flows. Therefore, the approach suggested by the authors and here proposed is to estimate a nonlinear model which accommodate for the two mentioned problems of OLS, that is the Poisson Pseudo-Maximum Likelihood estimator.

By construction the PPML implies that the dependent or response variable is given by the migration flows taken in levels, meaning that also zero observations can be included. As the economic theory suggests, these flows are supposed to be a positive function of the size of the economies involved and a negative function of the distance between countries. In particular, the size is usually represented with the population residing in the origin and destination countries, while the distance, specified as geographical distance between capital cities of the countries forming the dyad, accounts for migration costs. Since the link function of migration flows with the covariates is a log type, the measures of size and distance are taken in logarithmic terms. Following the approach of Lewer and Van den Berg (2008), the attractive forces between source and destination countries could be represented with the log of the product of population size in the source and host countries, moreover the attractiveness depends on the difference between the GDP per capita (always in log) in the two countries.

Although, it is worth noting that the simple inclusion of these variables ignores completely two important issues considered in the whole paper: MLR and spatial interdependence in gravity. In particular, with regard to MLR, we include origin-time and destination-time dummies, as suggested in the works of Ortega and Peri (2013) and Beine and Parsons (2015). Moreover, in order to take into account spatial interdependence between dyads, a diffusion effect is included and modeled through the econometric specification suggested in KMS (2014).

Summing up, for the additive error term ε_{ijt}, the specification of the gravity model for migration is described as follows:

$$Flows_{ijt} = \exp\Big(\ln(pop_{it} \cdot pop_{jt}) + \ln(gdp_{jt}/gdp_{it}) + \ln(dist_{ij}) + \ln(\widetilde{Flows}_{ijt}) + \zeta_{it} + \chi_{jt}\Big) + \varepsilon_{ijt}$$

where migration flows in levels between countries i and j at time t, $Flows_{ijt}$, depend on the product of source and destination countries' population in log terms, $\ln(pop_{it} \cdot pop_{jt})$, on the relative difference in GDP per capita between the target and origin country at time t, $\ln(gdp_{jt}/gdp_{it})$, the logarithm of geographical distance between capital cities of countries i and j (time-invariant variable),$\ln(dist_{ij})$, and the diffusion effect $\ln(\widetilde{Flows}_{ijt})$, denoting the spatial variable constructed through KMS and taken in logs. In addition, the model accounts for multilateral resistance with the origin-time fixed effects ζ_{it} and the target-time fixed effects χ_{ij}. No specific pair fixed effects are considered, otherwise the coefficient of distance cannot be identified.

As described in the section above, the analysis has been conducted on a sample including 16 ENC countries as source and 11 EU members as destination countries, identified on the basis of data availability. The model has been estimated with standard errors clustered for each origin and destination country combination, as suggested in the literature of PPML estimation, so accounting for both heteroskedasticity and autocorrelation.

Since for Syria, data on GDP per capita for the observed time-period are not available, we estimated the complete model for the sample excluding Syria and the model for the entire sample omitting the relative difference in income variable. The obtained results suggest an almost insignificant role played by the income term in determining the migration flows, so that we report in Tables 9-10 the results for the whole sample and in the Appendix (Tables 11-12) the results related to the sample without Syria.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(dist_{ij})$	-1.316***	-1.217**	-1.728***	-1.728***	-	-1.255***
$\ln(pop_{it} \cdot pop_{jt})$	-	0.763***	-	0.930***	-	-
$\ln(\widetilde{Flows_{ijt}})$	-	-	-	-	0.763***	0.607***
origin-time dummies	no	no	yes	yes	no	no
destination-time dummies	no	no	yes	yes	no	no

Table 2: Migration gravity model estimates with PPML
legend: * p<0.05; ** p<0.01; *** p<0.001

Variable	(7)	(8)	(9)	(10)	(11)	(12)
$\ln(dist_{ij})$	-1.186**	–1.691***	-1.164***	-	-1.658***	-1.658***
$\ln(pop_{it} \cdot pop_{jt})$	0.746***	-	-	-	-	0.258***
$\ln(\widetilde{Flows_{ijt}})$	0.373***	0.881***	0.819***	0.945***	0.743***	0.743***
origin-time dummies	no	yes	no	yes	yes	yes
destination-time dummies	no	no	yes	yes	yes	yes

Table 3: Migration gravity model estimates with PPML (continued)
legend: * p<0.05; ** p<0.01; *** p<0.001

The results are presented by gradually increasing the complexity of the model specification. The first four columns of Table 9 show the results of estimating the model without considering the spatial effect. In (1) and (2) we do not consider the presence of fixed effects and we proceed by estimating, in (1), the inverse relation between flows and distance and, in (2), we add the size of the considered economies measured with the log of the population product in source and destination countries. Models (3) and (4) add to the first two specifications the multilateral resistance issue by using origin-time and destination-time dummies. In specification (1) - (4), the coefficients relative to the mass of the source and destination economies are positive as expected but with respect to their magnitude, the introduction of origin-time and destination-time dummies produced a broader effect since the coefficient related to population size increases when fixed effects enter the equations. The same for the log of distance that affects significantly and negatively migration flows, in line with the Newton's gravity hypothesis which suggests an inverse relation between the two. In (5), we introduce the spatial variable capturing the so called diffusion effect without accounting for multilateral resistance, the same as in (6) and (7) where we gradually add the other regressors entering the traditional gravity equation. In regressions (8) and (9), we evaluate only the two determinants of migration flows related to distance, considering, alternately, origin-time dummies (8) and destination-time dummies (9). From (10) to (12), we consider both fixed effect and the spatial variable, going from the univariate case in (10) to the complete model in (12). In all model specifications, the introduction of the diffusion effect, related to the influence caused only by those country-pairs similar between them, has a significant and positive impact on migration. Specifically, looking at model (12), the diffusion effect mitigates the impact of countries' size, giving some evidence of the important role played by the socio-economic partnership between ENC and EU. Empirical insights suggest that the cultural, political and economic cooperation between different countries incentives people to move from home to host countries, by looking at the global actors of migration policy. In particular, with special regard to the spatial variable modeling, we can argue that migration flows within a specific country-pair are influenced by those between other source and target countries which share a similar distance with the dyad taken as reference point.

4. Final Remarks

In this work, we propose new empirical evidence in the estimation of gravity models for migration. We point out two fundamental issues which arise in the estimation procedure. First, we consider the drawback associated with the traditional OLS estimation of log-linearized models. In the presence of heterosckedasticity, the least squares estimates are seriously biased since the transformed error term is correlated with the considered covariates. Furthermore, the log transformation of migration flows is incompatible with the presence of zero values in the response variable, leading to the elimination of those data from the sample. In order to address these estimation problems, the proposed solution is to adopt an estimator which provides correct and consistent results. So that, following the suggestion of Santos Silva and Tenreyro (2006), the PPML estimator has been implemented to deal with these sources of biases. In particular, the advantages of PPML estimation are related to its flexible structure: for consistency, the correct specification of the conditional mean of migration flows, that is the choice of the right set of covariates, is a necessary condition. Additionally, zeros are not deleted from the sample, since migration flows enters in levels by construction.

Second, a widespread phenomenon which can lead to misleading results, if not considered, concerns the concept of multilateral resistance and spatial interdependence. In this work, we provide a new way to handle it, by adding to the usual inclusion of origin and destination-time dummies, a spatial variable able to capture a contagion effect in the dyadic setting of country-pairs relations. Specifically, we propose an adaptation of the spatial econometric modeling proposed by KMS (2014) in the field of nonlinear panel data. The spatial variable captures the aggregate influence of third part with respect to a specific country-pair, considering the average behaviour in migration policy of a group of “similar” country-pairs, through the implementation of a threshold mechanism on distances.

Finally, we apply this estimation methodology to a panel dataset formed by European neighbours and European members over a period of 18 years. The results confirm the positive influence of pull factors (such as the population size of origin and destination countries) and the negative impact of geographical distance on migration flows. But an important determinant in the volume of migration flows is given by the significant and positive contribution of the spatial term, meaning that the action of similar country-pairs produces an aggregate diffusion effect.

Appendix

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(dist_{ij})$	-1.673***	-1.625***	-1.691***	-1.583***	-1.583***	-
$\ln(pop_{it} \cdot pop_{jt})$	-	0.819***	0.856***	-	1.266***	-
$\ln(gdp_{jt}/gdp_{it})$	-0.258	-	0.353	0.362	0.629	-
$\ln(\widetilde{Flows_{ijt}})$	-	-	-	-	-	0.950***
origin-time dummies	no	no	no	yes	yes	no
destination-time dummies	no	no	no	yes	yes	no

Table 4: Migration gravity model estimates with PPML on the incomplete sample
legend: * p<0.05; ** p<0.01; *** p<0.001

Variable	(7)	(8)	(9)	(10)	(11)	(12)
$\ln(dist_{ij})$	-1.592***	-1.595***	-1.588***	-1.510***	-1.510***	-1.510***
$\ln(pop_{it} \cdot pop_{jt})$	-	-	0.801***	-	0.853***	-
$\ln(gdp_{jt}/gdp_{it})$	-	-0.253	-	0.376	-0.133	-
$\ln(\widetilde{Flows_{ijt}})$	0.672***	0.686***	0.435***	0.758***	0.758***	0.758***
origin-time dummies	no	no	no	yes	yes	yes
destination-time dummies	no	no	no	yes	yes	yes

Table 5: Migration gravity model estimates with PPML on the incomplete sample (continued)
legend: * p<0.05; ** p<0.01; *** p<0.001

References

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

Anderson, J. E. & Van Wincoop, E. (2003), 'Gravity with gravitas: a solution to the border puzzle', *The American Economic Review* **93**(1), 170–192.

Anderson, J. E. & Van Wincoop, E. (2004), 'Trade costs', *Journal of Economic Literature* **42**(3), 691–751.

Bai, J. (2009), 'Panel data models with interactive fixed effects', *Econometrica* **77**(4), 1229–1279.

Behrens, K., Ertur, C. & Koch, W. (2012), 'Dual gravity: Using spatial econometrics to control for multilateral resistance', *Journal of Applied Econometrics* **27**(5), 773–794.

Beine, M., Bertoli, S. & Fernández-Huertas Moraga, J. (2016), 'A practitioners guide to gravity models of international migration', *The World Economy* **39**(4), 496–512.

Beine, M. & Parsons, C. (2015), 'Climatic factors as determinants of international migration', *The Scandinavian Journal of Economics* **117**(2), 723–767.

Bergstrand, J. H. (1985), 'The gravity equation in international trade: some microeconomic foundations and empirical evidence', *The Review of Economics and Statistics* pp. 474–481.

Bertoli, S. & Moraga, J. F.-H. (2013), 'Multilateral resistance to migration', *Journal of Development Economics* **102**, 79–100.

Deardorff, A. (1998), Determinants of bilateral trade: does gravity work in a neoclassical world?, *in* 'The regionalization of the world economy', University of Chicago Press, pp. 7–32.

Egger, P. H. & Staub, K. E. (2016), 'Glm estimation of trade gravity models with fixed effects', *Empirical Economics* **50**(1), 137–175.

Elkins, Z., Guzman, A. T. & Simmons, B. A. (2006), 'Competing for capital: The diffusion of bilateral investment treaties, 1960–2000', *International Organization* **60**(4), 811–846.

Elkins, Z. & Simmons, B. (2005), 'On waves, clusters, and diffusion: A conceptual framework', *The Annals of the American Academy of Political and Social Science* **598**(1), 33–51.

Fahrmeir, L. & Tutz, G. (2013), *Multivariate statistical modelling based on generalized linear models*, Springer Science & Business Media.

Frankel, J. A., Stein, E. & Wei, S.-J. (1993), Continental trading blocs: are they natural, or super-natural?, Technical report, National Bureau of Economic Research.

Gourieroux, C., Monfort, A. & Trognon, A. (1984b), 'Pseudo maximum likelihood methods: applications to poisson models', *Econometrica: Journal of the Econometric Society* pp. 701–720.

Greene, W. H. (2003), *Econometric analysis*, Pearson Education India.

Grogger, J. & Hanson, G. H. (2011), 'Income maximization and the selection and sorting of international migrants', *Journal of Development Economics* **95**(1), 42–57.

Gunnella, V., Mastromarco, C., Serlenga, L. & Shin, Y. (2015), 'The euro effects on intra-eu trade flows and balance: Evidence from the cross sectionally dependent panel gravity models'.

Henderson, D. J. & Millimet, D. L. (2008), 'Is gravity linear?', *Journal of Applied Econometrics* **23**(2), 137–172.

Kapetanios, G., Mitchell, J. & Shin, Y. (2014), 'A nonlinear panel data model of cross-sectional dependence', *Journal of Econometrics* **179**(2), 134–157.

King, G. (2001), 'Proper nouns and methodological propriety: Pooling dyads in international relations data', *International Organization* **55**(2), 497–507.

LeSage, J. P. & Pace, R. K. (2008), 'Spatial econometric modeling of origin-destination flows', *Journal of Regional Science* **48**(5), 941–967.

Lewer, J. J. & Van den Berg, H. (2008), 'A gravity model of immigration', *Economics Letters* **99**(1), 164–167.

Luce, R. (1959), 'Individual choice behavior', *Wiley, New York* .

Mastromarco, C., Serlenga, L. & Shin, Y. (2016), Multilateral resistance and the euro effects on trade flows, *in* 'Spatial Econometric Interaction Modelling', Springer, pp. 253–278.

McFadden, D. (1974), ‘Analysis of qualitative choice behavior. zarembka, p.(ed.): Frontiers in econometrics’.

Neumayer, E. & Plümper, T. (2010a), ‘Spatial effects in dyadic data’, *International Organization* **64**(1), 145–166.

O’Hara, M. & Parmeter, C. F. (2013), ‘Nonparametric generalized least squares in applied regression analysis’, *Pacific Economic Review* **18**(4), 456–474.

Ortega, F. & Peri, G. (2013), ‘The effect of income and immigration policies on international migration’, *Migration Studies* **1**(1), 47–74.

Patuelli, R., Mussoni, M. & Candela, G. (2013), ‘The effects of world heritage sites on domestic tourism: a spatial interaction model for italy’, *Journal of Geographical Systems* **15**(3), 369–402.

Pesaran, M. H. (2006), ‘Estimation and inference in large heterogeneous panels with a multifactor error structure’, *Econometrica* **74**(4), 967–1012.

Plümper, T. & Neumayer, E. (2010), ‘Model specification in the analysis of spatial dependence’, *European Journal of Political Research* **49**(3), 418–442.

Ramos, R. (2017), Modelling migration, *in* ‘The Econometrics of Multi-dimensional Panels’, Springer, pp. 377–395.

Ramos, R. & Suriñach, J. (2017), ‘A gravity model of migration between the enc and the eu’, *Tijdschrift Voor Economische en Sociale Geografie* **108**(1), 21–35.

Ravenstein, E. G. (1889), ‘The laws of migration’, *Journal of the Royal Statistical Society* **52**(2), 241–305.

Robinson, P. M. (1987), ‘Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form’, *Econometrica: Journal of the Econometric Society* pp. 875–891.

Serlenga, L. & Shin, Y. (2007), ‘Gravity models of intra-eu trade: application of the ccep-ht estimation in heterogeneous panels with unobserved common time-specific factors’, *Journal of Applied Econometrics* **22**(2), 361–381.

Serlenga, L. & Shin, Y. (2013), ‘The euro effect on intra-eu trade: evidence from the cross sectionally dependent panel gravity models’, *Unpublished mimeo* pp. 1–30.

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

Small, K. A. & Rosen, H. S. (1981), 'Applied welfare economics with discrete choice models', *Econometrica (pre-1986)* **49**(1), 105.

Tenreyro, S. (2007), 'On the trade impact of nominal exchange rate volatility', *Journal of Development Economics* **82**(2), 485–508.

Tinbergen, J. (1962), 'An analysis of world trade flows', *Shaping the World Economy* **3**, 1–117.

Tobler, W. R. (1970), 'A computer movie simulating urban growth in the detroit region', *Economic Geography* **46**(sup1), 234–240.