Spatial Gravity

Michael Beenstock
Department of Economics
Hebrew University of Jerusalem

Daniel Felsenstein
Department of Geography
Hebrew University of Jerusalem

November 26, 2012

1. Introduction

Gravity modeling has been used extensively to investigate the determinants of economic relationships in trade, capital investment and migration between origins and destinations. The gravity model hypothesizes that bilateral flows depend on push-factors in the origins and pull-factors in the destinations, as well as measures of distance between origins and destinations. For example, migration from country A to country B depends on push-factors in A and pull-factors in B. In one-way gravity models the origins and destinations are different spatial units, such as A and B. In two-way gravity models a spatial unit is both an origin and a destination; e.g. there is migration from B to A as well as from A to B.

For the most part gravity models have been estimated under the assumption that bilateral flows are independent[1]. Specifically, it is assumed that the residual errors of bilateral flows are independent. For example, migration from A to B is independent of migration from A to C and from D to B. Since B and C are alternative destinations for migrants from A, it is unlikely that AB flows will be independent of AC flows. Also, it is unlikely that AB flows will be independent of DB flows unless A and D are independent. If A and D are geographically related, then AB and DB flows are unlikely to be independent. Ignoring the dependence between bilateral flows induced statistical inefficiency at best, and bias and inconsistency at worst. The latter arises in nonlinear models e.g. when zero bilateral flows, which are usually quite frequent in the data, are modeled as a probit or Poisson process[2].

The econometric shortcomings of assuming bilateral independence in estimating gravity models were first pointed out by LeSage and Pace (2008). They investigated these shortcomings in a two-way context (migration within the United States) in which dependence between bilateral flows was modeled as a spatial lag process. This assumes that dependence between bilateral flows results from geographic proximity. In the above example B and C are geographically close. They estimated separate spatial lag coefficients for origins and destinations. The dependence between bilateral flows does not have to be spatial; dependence

[1] Originally suggested by Pöyhönen (1963) the gravity model has become a workhorse in the empirical analysis of bilateral flows.

[2] For example, Helpman, Melitz and Rubinstein (2008) have suggested nonlinear estimators to treat zero bilateral flows. However, by assuming bilateral independence, their solution may introduce new biases. Burger, van Oort and Linders (2009 have suggested the use of a zero-inflated Poisson model.

might arise for other reasons. In the example above B and C do not have to be geographically close; they might be substitute destinations for immigrants even though they are far apart. This approach is taken by Behrens, Ertur and Koch (2012) in their gravity model for trade flows.

The econometric methodology in the present paper has much in common with LeSage and Pace (2008). However, there are a number of differences. First, we propose a lagrange multiplier test for spatial autocorrelation in OLS estimates of gravity models. This test is particularly useful because if the OLS residuals from gravity models are spatially independent there is unlikely to be a strong case for making the effort to estimate spatial lags in gravity models. Second, if the residuals from gravity models are spatially autocorrelated, the OLS standard errors are incorrect. We suggest a method for estimating robust standard errors for the parameters of gravity models estimated by OLS, which exploits the pattern of spatial autocorrelation between the origins and between the destinations. This extends Driscoll and Kraay (1998) who suggested a similar idea for spatial panel data. Third, we estimate the spatial counterpart to ARCH (or GARCH) models in time series. We refer to this phenomenon by SpARCH. In SpARCH models volatility is spatially related so that residual volatility depends upon residual volatility in neighboring regions (see for example Willocks 2010). In standard spatial models the conditional mean depends on its counterpart in neighboring spatial units. In SpARCH models both the conditional mean and its variance are spatially dependent. Fourth, the identities of origins and destinations are separate. Specifically, the origins are European Neighborhood Countries (ENCs) and the destinations are EU countries.

We illustrate these methods using migration data from the ENP countries to the EU during 2000 – 2010. The underlying hypothesis is that migration is motivated by economic differentials between the EU countries and the ENPs. Also, migration is motivated by the generosity of welfare benefits and the number of incumbent ENP migrants in the EU countries. Since the data include illegal immigrants, we also investigate whether immigration depends on the legal rights of illegal immigrants in the EU destinations.

Whereas the spatial econometrics of gravity models has been ignored, the same does not apply to the econometric treatment of zero bilateral flows (Helpman, Melitz and Rubinstein 2008, Burger, van Oort and Linders 2009). This problem arises from the fact that typically many bilateral flows are zero, in which case OLS is inappropriate. Various approaches have been suggested for treating these zeros, including Heckman type selection, tobit censoring, and zero-inflated count data analysis. There are zero bilateral flows in our ENP – EU migration data too. However, we side-step this issue by noting that these zeros occur exclusively among the newer members of the EU, and do not arise among the EU15. Therefore, our main econometric concern is with the spatial econometrics of one-way gravity models.

In section 2 we discuss the spatial econometrics of one-way gravity models. In section 3 we specify a theoretical model for one-way international immigration, specifically immigration from the ENPs to the EU15. An innovative component is the focus on welfare-induced immigration and policy towards illegal immigrants in the EU15. The data are presented in section 4 and results are reported in section 5. Section 6 concludes.

2. Methodology

Let Y_{od} denote a bilateral flow of Y from origin o to destination d, when there are N_o origins and N_d destinations. Therefore, the total number of observations is $N = N_oN_d$. Origins are labeled by $o = 1,2,..,N_o$ and destinations are labeled by $d = 1,2,\ldots,N_d$. Let j denote a neighbor of o and k denote a neighbor of d. The gravity hypothesis of interest is:

$$Y_{od} = \alpha + X_o\beta + Z_d\gamma + V\psi + u_{od} \qquad (1)$$

where X are push factors in the origin, Z are pull factors in the destination, and V is a vector of distance measures between origins and destinations. If equation (1) is estimated by OLS it is assumed by default that the residuals u_{od} are uncorrelated. However, as mentioned above this is unlikely for three main reasons. First, there may be intra-destination spatial autocorrelation so that $E(u_{od}u_{ok}) \neq 0$ Secondly, there may be intra-origin spatial autocorrelation so that $E(u_{od}u_{jd}) \neq 0$. Third, there may be spatial autocorrelation between origins and destinations so that $E(u_{od}u_{jk}) \neq 0$. In what follows we focus on intra-origin and intra-destination spatial autocorrelation because it is difficult to motivate spatial autocorrelation between origins and destinations.

Origins and destinations have different numbers of neighbors depending on their geography. Origin o has J_o neighbors labeled by j and destination d has K_d neighbors labeled by k. We use tildes to denote spatial variables. Equation (2a) defines the intra-origin spatial residual for origin o to destination d. It is a weighted average of the residuals of o's neighbors to destination d. The spatial intra-origin weights w_{oj} are summed to one for convenience.

$$\tilde{u}_{od} = \sum_{j=1}^{J_o} w_{oj} u_{jd} \qquad (2a)$$

$$\tilde{u}_{do} = \sum_{k=1}^{K_d} v_{dk} u_{ok} \qquad (2b)$$

Equation (2b) defines the intra-destination spatial residual for unit o in destination d. It is a weighted average of o's residuals in the neighbors of destination d. The intra destination weights v_{dk} are summed to one for convenience. However, these normalizations are not essential.

2.1 Spatial Autocorrelation

The lagrange multiplier approach to misspecification testing conveniently assumes that various restrictions do not apply, and tests the validity of these assumptions ex post. By contrast, likelihood ratio tests involve the estimation of both the restricted and the unrestricted models, and Wald tests require the estimation of the restricted model. Since the restricted model may be difficult to estimate, the LM approach has obvious practical advantages. The main disadvantage is that its statistical power is inferior. In the present context the restricted model would involve accounting for the dependence between bilateral flows in the gravity model, which is a difficult task in its own right. By contrast, the LM approach ignores these restrictions, estimates the gravity model by OLS and then checks the OLS residuals for spatial autocorrelation. If the LM statistic is not statistically significant, the OLS assumption of independence between bilateral flows is vindicated.

LM tests have to control for the covariates used to estimate the residuals, i.e. X_o and Z_d in equation (1). Therefore, the auxiliary regression for the LM test of SACD and SACO is given by equation (3):

$$\hat{u}_{od} = \phi + X_o\lambda + Z_d\mu + \theta_o\tilde{u}_{od} + \theta_d\tilde{u}_{do} + \varepsilon_{od} \qquad (3)$$

The intra-origin and intra-destination SAC coefficients are θ_o and θ_d respectively. The former picks up SAC among the origin's J_o neighbors. The latter picks up SAC among the destination's K_d neighbors. If $\theta_o = \theta_d = 0$ it must be the case that the elements of λ and μ are zero too since OLS assumes that X_o and Z_d are independent of the residuals. Therefore, if there is no SAC the R^2 of equation (3) must be zero. The

LM statistic is equal to NR^2 and it has a chi-square distribution with 2 degrees of freedom, one for intra-origin SAC and the other for intra-destination SAC. If the LM statistic exceeds its critical value it must be because θ_o and θ_d differ from zero.

The origin weights (w) and the destination weights (v) may be specified in terms of contiguity etc. We define them by:

$$w_{oj} = \frac{m_{oj}}{m_o} \qquad (4a)$$

$$v_{dk} = \frac{m_{dk}}{m_d} \qquad (4b)$$

where m_{oj} is the bilateral flow from o to neighbor j (intra ENC) and m_o is the total flow from o to its J_o neighbors, and m_{dk} is the flow from d to neighbor k (intra EU) and m_d is the total flow from d to its K_d neighbors. Because these weights are constructed out of intra-destination flows and intra-origin flows they are not directly dependent on flows from origins to destinations. For example, the inter o – d flow from Algeria to Belgium is independent of the intra o flow from Algeria to Morocco and the intra d flow from Belgium to Germany.

2.2 Robust Standard Errors

Spatial autocorrelation may be inherent or it might be induced by the misspecification of equation (1). In the latter case the remedy involves specifying the model correctly. In the former case the parameter estimates are unbiased but inefficient. In certain spatial panel data models they are also biased and inconsistent. Driscoll and Kraay (1998) suggested calculating “robust” standard errors, which take account of SAC in spatial panel data. We apply their approach to SAC in gravity models using the spatial covariance matrix for the residuals of equation (1).

Vectorizing equation (1) we rewrite it as:

$$y = Q\omega + u \qquad (5)$$

$$Q = (X\ Z)$$

$$\omega' = (\beta\ \gamma)$$

$$u = (\theta_o W + \theta_d V)u + \varepsilon \quad (6)$$

where W and V are the NxN spatial connectivity matrices with elements w and v respectively. The solution to equation (6) is:

$$u = A\varepsilon \qquad (7)$$

$$A = (I_N - \theta_o W - \theta_d V)^{-1}$$

The spatially robust covariance matrix of the OLS estimate of ω is:

$$\Sigma_\omega = (Q'Q)^{-1}(Q'\Theta Q)(Q'Q)^{-1} \qquad (8)$$
$$\Theta = A\Sigma_\varepsilon A'$$

If ε is homoscedastic $\Theta = \sigma_\varepsilon^2 AA$`. To implement equation (8) estimates of A and $\Sigma_\varepsilon$ based on estimates of $\theta_o$, $\theta_d$ and ε obtained from equation (3) are substituted into equation (8). If ε is heteroskedastic Θ = AΞA` where Ξ is a diagonal matrix with diagonal elements $\hat{u}_{od}^2$.

An obvious and asymptotically superior alternative to the use of spatially robust standard errors is to estimate equation (1) by FGLS, which involves the joint estimation of the parameters in equation (1) and θ_o and θ_d.

2.3 Spatially Autoregressive Conditional Heteroscedasticity (SpARCH)

The ARCH model developed for time series assumes that current volatility, as measured by the variance, depends on lagged volatility. In the case of residual volatility, the first-order ARCH model relates the squared residual at time t (u_t^2 which represents current volatility) to the squared residual at time t-1 (u_{t-1}^2):

$$u_t^2 = a + bu_{t-1}^2 \qquad (9)$$

where b is the ARCH coefficient. Conditional volatility is $E(u_t^2 / u_{t-1}) = a + bu_{t-1}^2$, which is heteroskedastic because it depends on t. However, unconditional volatility is homoskedastic since it equals a/(1-b), which does not depend on t. Since the homoskedasticity assumption of OLS refers to unconditional homoskedasticity ARCH has no implications for statistical inference. Matters are, however, different in nonlinear models. The LM test for (p-order) ARCH is carried out by substituting estimates of the OLS residuals in equation (9) and estimating a and b by OLS with T observations. The LM statistic is TR^2 and has a chi-square distribution with p degrees of freedom where R^2 refers to equation (9).

The spatial counterpart of equation (9) is:

$$u_i^2 = a + b\tilde{u}_i^2 \qquad (9)$$

according to which $E(u_i^2 / \tilde{u}_i) = a + b\tilde{u}_i^2$, i.e. the variance is conditionally heteroskedastic since it depends on i. To obtain the unconditional variance, equation (9) is vectorized:

$$u^2 = a + bWu^2 \qquad (10)$$

from which the unconditional variance is $a(I - bW)^{-1}$ which is homoscedastic because it does not depend on i.

In one-way gravity models the counterpart to equation (9) is:

$$u_{od}^2 = a + b_o \tilde{u}_{od}^2 + b_d \tilde{u}_{do}^2 \qquad (11)$$

where b_o is the SpARCH coefficient induced by volatility among the origins and b_d is its counterpart among the destinations. The unconditional variance is $a(I - b_oW - b_dV)^{-1}$. The LM test statistic is NR^2 which has a chi-squared distribution with 2 degrees of freedom.

2.4 Spatial Dynamics

If there is a common factor, spatial autocorrelation indicates that the spatial dynamics of model are miss-specified, and that a spatially static model with SAC is inferior to a spatially dynamic model without SAC[3]. In this case the specification of a first order lagged spatial dependent variable is the appropriate methodological response to SAC. However, if the residuals of a spatially static model are not spatially autocorrelated, this does not necessarily mean that the specification of spatial lagged dependent variables is inappropriate. We test for SAC and specify spatial lagged dependent variables in the origins and destinations, as well as spatial lags on the gravity variables (X and Z):

$$Y_{od} = \alpha + X_o\beta + Z_d\gamma + \tilde{X}_o\psi + \tilde{Z}_d\eta + \pi_o\tilde{Y}_o + \pi_d\tilde{Y}_d + u_{od} \qquad (12)$$

$$\tilde{X}_{no} = \sum_{j=1}^{J_o} w_{oj}X_{nj} \qquad \tilde{Z}_{nd} = \sum_{k=1}^{K_d} v_{dk}Z_{nk} \qquad \tilde{Y}_o = \sum_{j=1}^{J_o} w_{oj}Y_{oj} \qquad \tilde{Y}_d = \sum_{k=1}^{K_d} v_{dk}Y_{ok}$$

where π_o and π_d are the spatial lag coefficients, and ψ and η are the spatial lag coefficients on the push and pull variables.

LeSage and Pace (2008) and LeSage and Fischer (2010) discuss a variant of equation (12) in which flows are two-way so that all spatial units are both origins and destinations. This simplifies the specification of the spatial weights because there is no need to distinguish between origins and destinations. On the other hand, it greatly increases the burden of estimation because there are N^2 flows instead of N. In what follows we discuss modifications to the LeSage & Pace methodology when origins and destinations have different identities.

The observations are stacked (as in panel data) by flows from origin o to all N_d destinations. The flows from origin o are denoted by y_o which is a vector of length N_d. There are N_o such vectors. Define $y` = (y_1, y_2,\ldots,y_{No})`$, which is of length N_dN_o. The N_d x N_d spatial weights matrix among the destinations with elements v_{de} is denoted by

[3] See e.g. Anselin (1988) pp 226-9.

W_d and the N_o x N_o spatial weights matrix among the origins with elements w_{op} is denoted by W_o. W_d and W_o have zeros on their leading diagonal. We may write y as:

$$\tilde{y} = Dy + \Omega y = \tilde{y}_d + \tilde{y}_o \qquad (13)$$

$$D = I_{No} \otimes W_d \qquad \Omega = W_0 \otimes I_{Nd}$$

Equation (13) decomposes the spatial lagged dependent variable into its destination and origin components. Ω and D are N x N matrices ($N = N_oN_d$). D is block diagonal with W_d on the leading diagonal and N_d x N_d zero matrices elsewhere. Ω is made up of N_o blocs of matrices $w_{op}I_{Nd.}$

Equation (12) may be expressed in terms of these matrices:

$$y = \alpha + X\beta + Z\gamma + \Omega X\psi + DZ\eta + \pi_o\Omega y + \pi_d Dy + u \qquad (14)$$

The only difference between equation (14) and standard spatial lag models is the separate terms in Ωy and Dy. If $\pi_o = \pi_d = \pi$ equation (14) simplifies to the standard spatial lag model[4]:

$$y = \alpha + X\beta + Z\gamma + \Omega X\psi + DZ\eta + \pi Wy + u \qquad (15)$$

$$W = D + \Omega$$

Because equation (15) contains a single spatial lag coefficient (π) it is standard, and it may be estimated by maximum likelihood using statistical packages such as Matlab. Matters are different for equation (14) because it contains two spatial lag coefficients (π_o and π_d), and the likelihood function depends on $\ln|I_N - \pi_o\Omega - \pi_d D|$. Elhorst et al (2012) have developed estimators for higher order spatial lag models, which are used below to estimate π_o and π_d.

According to equation (1) developments in destination k (d's neighbor) do not affect migration from o to d. Nor do developments in origin j (o's neighbor) affect migration from from o to d. Indeed, migration from o to d depends only on developments in d and o; the bilateral flows are independent. According to equation (14) these bilateral flows are dependent for two reasons. First, the spatial lag coefficients (η and ψ) on the push and pull factors mean that migration from o to d depends on developments in k and j. Second, the coefficients on the spatial lagged dependent variable (π_o and π_d) mean that migration from o to d depends on developments beyond the neighbors of o and d.

---

[4] This restriction is imposed using Matlab by Behrens et al (2012).

If for simplicity $\alpha = \eta = \psi = 0$ and there is one push factor z and pull factor x, the solution to equation (14) is:

$$y = C(\beta x + \gamma z) \qquad (16)$$

$$C = \left(I_N - \pi_o \Omega - \pi_d D\right)^{-1}$$

Therefore:

$$Y_{od} = \beta \sum_{n=1}^{N} c_{od.n} x_n + \gamma \sum_{n=1}^{N} c_{od.n} z_n \qquad (17)$$

where $c_{od.n}$ are the elements of C from row od. Equation (17) shows that migration from o to d depends on x and z in all origins and destinations. However, these elements are likely to vary inversely with their distance from o and d. If $\pi_o = \pi_d = 0$ equation (17) simplifies to:

$$Y_{od} = \beta x_o + \gamma z_d \qquad (18)$$

i.e. only developments in o and d affect migration from o to d. Notice that the coefficients of x_o and z_d in equation (17) differ from their counterpart in equation (18) since $c_{od.od}$ generally differs from 1. Indeed, positive spatial feedback implies that $c_{od.od}$ is likely to exceed 1.

2. Immigration Theory

The basic idea that immigration is driven by income differentials between origins and destinations is usually attributed to Hicks (1932) and Sjaastad (1962). However, Adam Smith argued that migration is driven by wage differentials, and regarded policies to limit internal migration in England immigration as unjust and economically harmful[5]. The development of the welfare state during the 20th century created a new motivation for immigration. Immigrants are attracted to destinations where welfare benefits in cash and in kind are more generous[6]. Empirical evidence in favor of this hypothesis has been found for the EU (Péridy 2006, De Giorgi and Pellizzari 2006, Docquier et al 2006 and Razin et al 2011) and for internal migration in the US (Borjas 1999, McKinnish 2007). Razin et al argue that welfare generosity disproportionately attracts unskilled immigrants because skilled immigrants are

[5] Smith (1976) argued that the law of settlements, enacted to enforce poor law benefits provide by parishes, restricted internal migration and were responsible for spatial wage inequality. “The very unequal price of labour which we frequently find in England in places at no great distance from one another, is probably owing to the obstruction which the law of settlements gives to a poor man who would carry his industry from one parish to another without a certificate.” (p 142). Smith called for the repeal of the law of settlements and the promotion of economically motivated migration.

[6] Adam Smith would have been familiar with this theory since the law of settlements prevented individuals from migrating to parishes where the poor laws were administered more generously.

deterred by the higher taxation required to finance this generosity. In all of these studies it is assumed that bilateral migration flows are independent.

2.1 Stocks and Flows

Immigration flows during time t to t+1 are hypothesized to be determined according to Sjaastad's stock adjustment model in which the levels of push and pull factors at time t and their changes during times t to t+1 are hypothesized to determine immigration flows from origins to destinations. For example, if GDP per head is a pull factor in the destinations, immigration varies directly with the level of GDP per head at time t and the change in GDP per head between times t and t+1. If the latter happens to be zero, immigration depends only on the initial level. If the immigrant stock was at its equilibrium level in time t, the stock-adjustment model predicts that immigration during times t and t+1 should be zero.

The stock adjustment model should control for the stock of immigrants at time t. Given everything else the effect of the initial stock should be negative. If, however, incumbent immigrants provide new immigrants with social network amenities, the stock of immigrants at time t might also increase immigration (isn't there a cut off point beyond which the size of stock has no effect?).

Let y_{odt} denote the stock of immigrants from o in d in time t and y^*_{odt} denote its equilibrium counterpart. The stock adjustment model predicts that the flow of immigrants between times t and t+1 is:

$$Y_{odt} = \phi(y^*_{odt} - y_{odt}) + \varphi\Delta y^*_{odt+1} \qquad (19)$$

where ϕ and φ are stock adjustment coefficients. Let P_d denote a vector of p pull factors in d, let U_o denote a vector of push factors in o, and let P denote the N_d x p matrix with rows P_d. In principle, immigrants from o may choose between all destinations. Therefore:

$$y^*_{odt} = \Gamma_o P_t + \Xi_o U_{ot} \qquad (20)$$

Equation (20) states that the equilibrium number of immigrants from o in d at time t depends via Γ_o on the pull factors in d and rival destinations, as well as the push factors in o. Substituting equation (20) into equation (19) and assuming Γ and Ξ do not vary by origin gives:

$$Y_{odt} = \phi(\Gamma P_t + \Xi U_{ot}) + \varphi(\Gamma\Delta P_{t+1} + \Xi\Delta U_{ot+1}) - \phi y_{odt} \qquad (21)$$

Therefore in equation (1) $X_o = P_t + \Delta P_{t+1}$ and $Z_o = U_{ot} + \Delta U_{ot+1}$. Equation (21) is a multilateral gravity model because bilateral flows depend on multilateral nodes.

Tunisians may emigrate to France as well as other EU countries. According to equation (21) they compare pull factor in France with pull factors in other EU countries (giving appropriate spatial weights).

One of these pull factors may be the existing number of Tunisians in France relative to other EU countries. Therefore, y_{odt} may be a pull factor. If so, this variable has a positive effect as a pull factor, and a negative effect as indicated in equation (21).

2.2 Push and Pull Factors

In gravity models immigration is assumed to depend on GDP per head in origins and destinations, as well as measures of cultural and ethnic difference. For example, if o and d share a common language immigration from o to d is likely to be greater. Also, immigration is hypothesized to vary inversely with the geographical distance between o and d. If immigrants are positively selected (Borjas 1987) they are attracted by income inequality since they expect to earn more where there is more dispersion. If so, immigration should vary directly with the gini coefficient in d.

We also investigate whether immigration is motivated by welfare. Legal immigrants benefit from social security and other benefits received by natives. Apart from pecuniary benefits, such as unemployment benefit and income support, we attach importance to benefits in kind including health, education and housing. Given everything else, we expect that d will be a more attractive destination to immigrants the more generous are its benefits.

The case of illegal (the politically correct EU parlance is ‘ irregular’) immigrants is more complicated. Procedures for dealing with political refugees vary by country; they may be more or less lenient. If d is more lenient it is likely to attract more immigrants. Illegal immigrants either did not apply for refugee status, or if they did and were refused, they go underground. Countries also vary by their alacrity in expelling illegal immigrants. Finally, countries vary by the legal rights of illegal immigrants and their children in terms of their access to health services and schooling. Countries that are more lenient and generous in their treatment of illegal immigrants are expected to be more attractive as destinations. We are unaware of empirical studies of the effects of immigration policy on illegal immigration. Indeed, Yoshida and Woodland (2005) signally do not cite such studies[7].

3. Data

[7] Their concern is with the effects of illegal immigration on natives and policies designed to achieve the socially optimum amount of illegal immigration.

We use data from the Global Bilateral Migration Database (GBMD), compiled by the World Bank, on stocks and flows of immigrants from ENCs in EU countries. These data have been compiled for almost all countries of the world and are based on census data. The stock data refer to 1960, 1970,..,2010, and the flow data refer to decades e.g. 2000 – 2010. Flows are defined to equal changes in stocks. Therefore, return migration, for example, by Tunisians in France appropriately reduces the number of Tunisians in France in the data. If Tunisians in France migrate to third countries e.g. Belgium, the number of Tunisians in Belgium increases and the number in France decreases. In the data the flow of immigrants from Tunisia to France decreases and the flow from Tunisia to Belgium increases[8]. Table 1 reports the immigrant flow data for 2000 – 2010 from ENC origins to EU destinations, and Table 2 reports the immigrant stock data in 2000. For example, in 2000 there were 8004 immigrants from Algeria in Belgium, but this number grew by 13,546 by 2010.

We have collected data on the rights of legal and illegal immigrants, as well as on the way countries treat illegal immigrants. We use data on expulsions and apprehensions to calculate expulsion and apprehension rates (in terms of the population at risk) in EU destinations. These rates are of the order of one percent except in Greece where they approach 30 percent (see data appendix). We also report in the data appendix an index of the treatment of legal immigrants in EU destinations in terms of the assistance they get to integrate economically, socially and politically.

Results

The dependent variable in equation (1) is defined as the rate of immigration that took place between 2000 and 2010, i.e. it is the data in Table 1 divided by the data in Table 2. The origin variables (Z) include GDP per head in 2000 and its rate of growth during 2000 – 2010. The destination variables (X) include GDP per head in 2000 and its rate of growth during 2000 – 2010, the gini coefficient for household income, social spending per head in 2000 and its rate of growth during 2000 – 2010, spending per head on primary education, expulsion and apprehension rates, and the treatment index of immigrants. We also control for distances between origins and destinations, common official languages, and immigrant stocks in 2000.

[8] It is unclear how the World Bank tracks these immigrants. If a Tunisian in France emigrates to Belgium as a Frenchman, it is not clear how his Tunisian identity is obtained unless census data gives place of birth. Also censuses are not decennial or coordinated. An appendix on GBMD downloaded from the website of the World Bank is provided. Also, GBMD data for Israel are different to their counterparts published by Israel's Central Bureau of Statistics.

Most of these variables turned out to be not statistically significant. Model 1 in Table 3 retains the variables which survived a specification search process in which insignificant variables were successively omitted. Since Model 1 is estimated by OLS it is assumed that the observations are spatially independent. The signs of the parameters in Model 1 are "correct" but they are not statistically significant at conventional levels. Since the LM test statistic for heteroskedasticity is highly significant, we also use robust standard errors.

Variables that do not feature in Model 1 include GDP per head and its growth in the EU destinations as well as the treatment index of immigrants. Immigration flows vary inversely with apprehension rates, and GDP per head and its growth in the EPC origins, and they vary directly with social spending per head, spending on education and income inequality. When model 1 is estimated using data for 1990 – 2000 its explanatory power is even smaller than it is for 2000 – 2010, none of the estimated parameters is statistically significant, and many parameters change their signs. In short, model 1 is not robust and depends on the observation period.

The LM statistics reported in Table 4 indicate that the residuals of model 1 are not spatially autocorrelated, and the SpARCH coefficients are not significantly different from zero. When spatially lagged dependent variables are specified in models 2 and 3, the spatial lag coefficients are statistically significant. In model 2 the spatial lag coefficients are restricted to be identical in origins and destinations. Although in model 3 these coefficients are unrestricted, their estimates turn out to be similar, but different to their counterpart in model 2. Table 4 shows that when spatially lagged dependent variables are specified, the SAC and SpARCH coefficients are statistically significant.

Discussion

In this paper we tried to make two contributions, methodological and substantive. Standard econometric analysis of gravity models has typically assumed that the observations are independent. This assumption is surprising because it implies that flows from a given origin to alternative destinations are independent. It also assumes that flows from different origins to the same destination are independent. We suggest a lagrange multiplier statistic to test origin – destination independence. We also model origin – destination dependence using recently developed double spatial lag estimators.

Our substantive contribution uses data on migration flows from European Neighborhood countries to EU destinations during the first decade of the 20th century to test key hypotheses concerning the determinants of international migration. These include the hypotheses that migration is driven by income differentials, income inequality, welfare generosity in the destination countries, and policies to deter irregular immigration.

During the first decade of the 20th century there is little if any evidence that migration from European Neighborhood Countries to the European Union depended on determinants that have been high-lighted in the theoretical literature. Neither the level of GDP per head in EU countries nor its rate of growth, explain migration from EN to EU. Therefore, the recent economic recession in EU is unlikely to deter migration from EN. There is some weak evidence that GDP per head and its growth in the EN countries deter migration. There is also some evidence that migrants prefer to migrate to EU countries where there is greater economic inequality. If immigrants are positively selected they stand to gain more in countries where incomes are more unequal.

There is no evidence that immigrants engage in welfare-chasing. This is true when welfare generosity is measured by social spending per head in the EU countries, when it is measured by per capita spending on primary schooling, or when expert indices are used. Nor does physical distance or common languages, which are standard variables in gravity models, significantly explain immigration from EN to EU. Indeed, immigration does not seem to be explained by any of the standard hypotheses regarding international migration. However, there is weak evidence that immigration policy, as measured by apprehension rates among irregular immigrants, deters immigration.

These results may be disappointing as far as policy recommendations are concerned. On the other hand, the methodological results are more salient. They show that results obtained using conventional econometric methods which assume gravity flows are independent are over-turned when these flows are specified to be dependent. Specifically, gravity models in which spatial lags are specified produce different results to standard gravity models. Moreover, separate spatial lags are specified among destination countries in the EU and origin countries in the EN. The coefficients on these spatial lags are about 0.5 – 0.6, implying that there are strong spillover effects in migration between neighboring origins as well as destinations.

Indeed, these effects cancel out almost all the substantive effects to which reference has already been made.

Immigration flow 2000-2010

Destination / Origin	Austria	Belgium	Denmark	Finland	France	Germany	Greece	Ireland	Italy	Luxem- -bourg	Netherlands	Portugal	Spain	Sweden	UK
Algeria	188	13,546	255	313	-143,341	786	111	1,084	13,619	58	-40	267	40,077	559	-24,943
Armenia	-65	958	175	59	11,520	-6,180	1,331	83	274	1	1,660	59	9,886	440	774
Azerbaijan	56	98	63	28	-32	21,210	54	69	155	1	2,398	14	425	292	749
Belarus	181	496	321	102	287	25,321	205	538	3,866	9	444	181	3,162	713	1,500
Egypt	5,321	1,872	465	311	22,964	6,684	2,067	836	46,986	22	1,920	144	2,554	761	1,127
Georgia	327	162	48	27	-14,264	-56,940	19,840	216	998	2	828	94	9,361	211	656
Israel	471	2,125	581	356	3,800	4,728	416	366	471	16	892	137	2,060	685	5,802
Jordan	132	334	373	107	324	4,667	321	198	736	1	45	17	1,165	534	3,490
Lebanon	955	3,375	3,390	225	34,422	9,977	2,531	224	5,981	19	294	83	1,838	4,614	4,393
Libya	64	335	65	56	1,108	3,638	105	1,022	-1,409	3	125	16	1,287	183	11,972
Morocco	343	61,720	1,644	594	578,523	23,823	188	287	189,285	110	16,101	815	525,278	1,799	-8,388
Moldova	140	226	81	43	-1,881	3,689	1,887	2,377	82,508	3	130	1,340	15,718	168	429
Russia	3,788	31,550	2,450	7,007	-174,649	-679,197	21,133	4,042	13,122	87	-17,563	1,825	50,042	4,981	18,253
Syria	1,343	2,235	943	147	10,674	14,242	5,288	162	1,191	6	1,039	49	2,734	5,386	-125
Tunisia	1,038	7,366	249	223	-8,586	11,789	131	124	45,900	46	433	75	1,716	911	-5,882
Ukraine	1,742	1,433	5,136	585	3,465	144,338	11,754	3,462	158,816	42	1,373	5,592	69,788	1,473	24,196

Immigration stock 2000

Destination / Origin	Austria	Belgium	Denmark	Finland	France	Germany	Greece	Ireland	Italy	Luxem- -bourg	Netherlands	Portugal	Spain	Sweden	UK
Algeria	546	8,004	932	456	1,057,135	20,295	267	861	15,861	347	3,873	0	23,269	1,664	40,555
Armenia	654	195	569	89	2,961	21,695	7,438	52	280	6	252	19	2,502	448	15
Azerbaijan	140	13	125	41	382	2,055	102	43	99	4	423	2	144	249	2
Belarus	373	45	239	154	791	3,813	336	610	1,680	42	71	5	667	590	46
Egypt	6,661	724	1,247	388	5,060	14,208	7,156	620	43,477	107	9,381	0	1,631	2,062	26,975
Georgia	332	254	110	47	15,420	75,104	21,977	150	318	12	113	105	1,341	174	82
Israel	1,696	1,679	1,423	442	4,919	9,351	335	285	2,561	74	4,314	0	912	1,500	7,729
Jordan	412	289	961	133	635	11,007	646	137	2,983	6	827	0	1,202	1,056	636
Lebanon	544	1,016	11,982	283	11,033	51,611	1,228	151	4,163	92	3,060	0	1,657	19,817	11,219
Libya	357	61	167	68	413	831	188	737	3,382	15	466	0	438	370	136
Morocco	827	110,962	4,776	998	262,462	84,619	521	302	286,498	557	151,254	1,094	253,173	4,443	20,878
Moldova	308	135	109	65	2,608	13,736	5,492	958	6,680	15	22	2,947	1,833	97	180
Russia	4,895	1,129	2,669	10,527	217,690	978,793	16,847	2,695	14,864	461	23,439	1,462	11,316	8,579	15,053
Syria	825	690	1,328	183	5,550	26,114	5,334	153	3,411	33	5,662	0	2,720	14,005	5,646
Tunisia	1,710	3,762	728	292	310,949	25,260	225	125	75,808	237	3,800	0	1,005	2,698	9,948
Ukraine	2,534	540	1,056	878	11,687	58,163	13,082	1,566	13,755	204	225	9,843	18,491	1,919	783

Table 3 Estimates of the Migration Model: 2000-2010

<table>
<tr><th></th><th colspan="2">Model 1: OLS</th><th colspan="2">Model 2: ML</th><th colspan="2">Model 3: ML</th></tr>
<tr><th></th><th>Coefficient</th><th>t statistic</th><th>Coefficient</th><th>t statistic</th><th>Coefficient</th><th>t statistic</th></tr>
<tr><td>Intercept</td><td>-0.66</td><td>-0.58</td><td>-0.558</td><td>-0.56</td><td>-0.54</td><td>-0.53</td></tr>
<tr><td>Immigrant stock 2000*</td><td>0.013</td><td>1.53</td><td>0.0091</td><td>1.39</td><td>-0.000387</td><td>-0.06</td></tr>
<tr><td>GDP per head in origin 2000*</td><td>-0.0314</td><td>-1.31</td><td>-0.00174</td><td>0.08</td><td>-0.00373</td><td>-0.18</td></tr>
<tr><td>Growth of GDP per head in origin</td><td>-0.0137</td><td>-0.99</td><td>-0.00735</td><td>-0.61</td><td>-0.00292</td><td>-0.24</td></tr>
<tr><td>Gini</td><td>1.709</td><td>1.95</td><td>1.115</td><td>1.54</td><td>0.7435</td><td>0.99</td></tr>
<tr><td>Social spending per head*</td><td>0.3283</td><td>0.31</td><td>0.0243</td><td>0.25</td><td>0.00384</td><td>0.40</td></tr>
<tr><td>Spending per pupil in primary education</td><td>0.0111</td><td>1.65</td><td>0.00477</td><td>0.94</td><td>0.00422</td><td>0.83</td></tr>
<tr><td>Apprehension rate</td><td>-3.02</td><td>-1.22</td><td>0.1129</td><td>0.25</td><td>0.3263</td><td>0.71</td></tr>
<tr><td>Common language</td><td>0.141</td><td>1.76</td><td>0.0968</td><td>1.41</td><td>0.0393</td><td>0.57</td></tr>
<tr><td>Distance</td><td>-0.000035</td><td>-1.50</td><td>-0.0000376</td><td>-1.86</td><td>-0.0000179</td><td>-0.90</td></tr>
<tr><td>Spatial lag: origin</td><td></td><td></td><td rowspan="2">0.09897</td><td rowspan="2">2.4675</td><td>0.500119</td><td>13.85</td></tr>
<tr><td>Spatial lag: destination</td><td></td><td></td><td>0.569238</td><td>16.65</td></tr>
<tr><td>R^2 adj</td><td colspan="2">0.0632</td><td colspan="2">0.0592</td><td colspan="2">0.0677</td></tr>
</table>

Dependent variable is the rate (percent) of migration from ENC to EU during 2000 – 2010. Asterisked variables are in logarithms.

Table 4 SAC and SpARCH Coefficients

Model	1	2	3
SAC			
Origin	0.0504 (0.23)	-0.4768 (-2.06)	-0.9941 (-9.16)
Destination	-0.0511 (-0.63)	-0.0840 (-0.37)	-0.9725 (-8.90)
LM	2.6015	24.209	81.399
SpARCH			
Origin	0.6596 (0.59)	0.9152 (4.18)	0.5922 (4.33)
Destination	0.0167 (0.25)	0.2350 (2.44)	0.5961 (6.91)
LM	0.408	25.536	61.968

Notes: LM refers to lagrange multiplier statistics for SAC and SpARCH. Their critical values ($p = 0.05$) are χ^2 ($df = 2$) = 5.991. t-statistics for SAC and SpARCH coefficients reported in parentheses.

Appendix 1

Global Bilateral Migration Database (World Bank)

"Global matrices of bilateral migrant stocks spanning the period 1960-2000, disaggregated by gender and based primarily on the foreign-born concept are presented. Over one thousand census and population register records are combined to construct decennial matrices corresponding to the last five completed census rounds. For the first time, a comprehensive picture of bilateral global migration over the last half of the twentieth century emerges."
World Bank Website

The table below compares data from GBMD for Israel with data published by Israel's Central Bureau of Statistic (CBS). The discrepancies are relatively small until 1980, but become large by 2010. These discrepancies cast doubt on the reliability of GBMD for other countries.

Country of Origin	CBS	GBMD	CBS	GBMD	CBS	GBM
	1980	1980	1990	1990	2000	200
GRAND TOTAL	**1,447,400**	**1,428,869**	**1,503,700**	**1,621,978**	**1,957,793**	**2,231,**
Asia - total	303,400	302,190	271,400	255,882	235,695	265,2
Turkey	44,300	1,545	37,800	15,410	31,256	34,7
Iraq	104,100	105,136	89,600	86,214	76,830	85,6
Yemen	52,300	53,470	45,600	42,746	36,968	40,9
Iran	52,000	59,753	57,100	53,077	51,638	58,3
India and Pakistan	19,500	20,921	19,100	18,650	18,145	20,5
Syria and Lebanon					13,044	14,5
Other	31,200	61,365	22,200	39,785	7,815	10,4
Africa - total	336,500	337,130	320,800	314,853	320,102	365,5
Morocco	216,300	217,033	191,700	182,652	167,372	187,8
Algeria and Tunisia	51,600	55,348	47,400	46,375	42,265	47,4
Libya	27,200	1,431	23,700	9,889	19,632	21,8
Egypt	30,500	31,199	26,100	25,275	22,133	24,6
Ethiopia					56,308	66,9
Other	10,900	32,119	31,900	50,662	7,815	16,8
Europe, America and Oceania - total	807,600	789,549	911,500	1,051,243	1,401,996	1,600,
USSR (former)	206,100	272,571	365,700	579,468	907,209	1,065,
Poland	174,100	175,609	126,300	121,526	83,316	89,7
Romania	183,700	193,905	154,700	151,154	125,793	139,2
Bulgaria and Greece	35,100	37,187	28,800	28,848	23,739	26,3
Germany and Austria	45,300	49,513	38,800	39,086	32,920	36,7
Czech Republic, Slovakia and Hungary	46,700	49,839	39,000	37,084	28,671	31,6
France					27,389	31,9
United Kingdom					18,626	21,4
Europe, other	43,000	6,039	60,400	30,391	30,994	11,7
North America and Oceania	37,400	2,690	50,400	35,030	69,478	80,0
Argentina					31,672	36,9
Latin America, other	30,000	2,196	47,400	28,656	22,191	28,5

Appendix 2: Data Sources

Variable	Unit	Definition	Source	Link
Immigration stock	Persons	Stock of persons born in country A living in country B at time t	World Bank - Global Bilateral Migration Database	http://data.worldbank.org/data-catalog/global-bilateral-migration-database
Immigration flow	Persons	Stock of persons born in country A living in country B at time t minus stock of persons born in country A living in country B at time t-1	World Bank - Global Bilateral Migration Database	http://data.worldbank.org/data-catalog/global-bilateral-migration-database
GDP	U.S. Dollars, current prices	Gross domestic product per capita	IMF - World Economic Outlook Databases	http://www.imf.org/external/pubs/ft/weo/2012/02/weodata/download.aspx
Education expenditure	%	Public expenditure per pupil as a % of GDP per capita	UNESCO	http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=143&IF_Language=eng
Inequality	Gini coefficient		OECD	http://stats.oecd.org/
Social expenditure	U.S. Dollars, constant PPPs (2000)	Expenditure per head	OECD	http://stats.oecd.org/
Common language	-	Common official language	CEPII Geodist dyadic dataset	http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Distance	Km	Simple distance between most populated cities	CEPII Geodist dyadic dataset	http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Labour Market Mobility	Index	Experts index on the Labour Market Mobility of immigrants	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/mipexrawdata_final_13_02_2012.xlsx
Family Reunion	Index	Experts index on the possibility of family reunion of	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/

		immigrants		mipexrawdata_final_13_02_2012.xlsx
Education	Index	Experts index on the special attention given to immigrant s needs in the education system	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/mipexrawdata_final_13_02_2012.xlsx
Political Participation	Index	Experts index on the level of political participation of immigrants	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/mipexrawdata_final_13_02_2012.xlsx
Long Term Residence	Index	Experts index on the long term residency possibilities for immigrants	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/mipexrawdata_final_13_02_2012.xlsx
Access to Nationality	Index	Experts index on access to nationality possibilities for immigrants	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/mipexrawdata_final_13_02_2012.xlsx
Anti-Discrimination	Index	Experts index on anti-discrimination regulations to protect immigrants	MIPEX – Migrant Integration Policy Index	http://www.mipex.eu/sites/default/files/downloads/mipexrawdata_final_13_02_2012.xlsx
Toleration of residence	Index	Index based on policy options for persons not removed due to practical or technical obstacles	FRA (European Union Agency for Fundamental Rights) - Fundamental rights of migrants in an irregular situation in the European Union	http://research.icmpd.org/fileadmin/Research-Website/FRA/FRA_irregular_migration/Final_Reports-FRA_published_2011/FRA_2011_Migrants_in_an_irregular_situation_EN.pdf
Crime	Index	Index based on whether irregular entry/stay considered a crime?	FRA (European Union Agency for Fundamental Rights) - Fundamental rights of migrants in an irregular	http://research.icmpd.org/fileadmin/Research-Website/FRA/FRA_irregular_migr

			situation in the European Union	ation/Final_Reports-FRA_published_2011/FRA_2011_Migrants_in_an_irregular_situation_EN.pdf
Housing	Index	Index based on the level of punishment for renting shelter to migrants in an irregular situation	FRA (European Union Agency for Fundamental Rights) - Fundamental rights of migrants in an irregular situation in the European Union	http://research.icmpd.org/fileadmin/Research-Website/FRA/FRA_irregular_migration/Final_Reports-FRA_published_2011/FRA_2011_Migrants_in_an_irregular_situation_EN.pdf
Healthcare	Index	Index based on the general healthcare entitlements for migrants in an irregular situation	FRA (European Union Agency for Fundamental Rights) - Fundamental rights of migrants in an irregular situation in the European Union	http://research.icmpd.org/fileadmin/Research-Website/FRA/FRA_irregular_migration/Final_Reports-FRA_published_2011/FRA_2011_Migrants_in_an_irregular_situation_EN.pdf
Education	Index	Index based on the right to education for undocumented children	FRA (European Union Agency for Fundamental Rights) - Fundamental rights of migrants in an irregular situation in the European Union	http://research.icmpd.org/fileadmin/Research-Website/FRA/FRA_irregular_migration/Final_Reports-FRA_published_2011/FRA_2011_Migrants_in_an_irregular_situation_EN.pdf
Apprehensions	%	% of the number of foreign nationals apprehended/found to be illegally staying vs. the migrant stock in the destination country	EMN (European Migration Network) - Annual Report on Migration and International Protection Statistics 2003-2009	http://emn.intrasoft-intl.com/Downloads/prepareShowFiles.do?entryTitle=2%2E%20Annual%20Reports%2

				0on%20Migration%20and%20International%20Protection%20Statistics
Refusals	%	% of the number of foreign nationals refused entry vs. the migrant stock in the destination country	EMN (European Migration Network) - Annual Report on Migration and International Protection Statistics 2003-2009	http://emn.intrasoft-intl.com/Downloads/prepareShowFiles.do?entryTitle=2%2E%20Annual%20Reports%20on%20Migration%20and%20International%20Protection%20Statistics
Removed	%	% of the number of foreign nationals removed vs. the migrant stock in the destination country	EMN (European Migration Network) - Annual Report on Migration and International Protection Statistics 2003-2009	http://emn.intrasoft-intl.com/Downloads/prepareShowFiles.do?entryTitle=2%2E%20Annual%20Reports%20on%20Migration%20and%20International%20Protection%20Statistics

References

Anselin J-L. (1988) *Spatial Econometrics: Methods and Models*, Dordrecht, Kluwer.

Behrens K., C. Ertur and W. Koch (2012) "Dual gravity": using spatial econometrics to control for multilateral resistance. *Journal of Applied Econometrics*, 27: 773-794.

Borjas G.S. (1988) Self selection and the earnings of immigrants. *American Economic Review*, 77: 531-533.

Borjas G.S. (1999) *Heaven's Door: Immigration Policy and the American Economy*, Princeton University Press.

Burger M., van Oort F and Linders G-J (2009), On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation, *Spatial Economic Analysis*, 4 (2), 167-190

De Giorgi G. and M. Pellizzari (2006) Welfare migration in Europe and the cost of harmonized social assistance, IZA discussion paper 2094.

Docquier F. and A. Marfouk (2006) International migration by education attainment, 1990 – 2000. *International Migration, Remittances, and the Brain Drain*, C. Ozden and M. Schiff (eds), New York, Palgrave Macmillan.

Driscoll J.C. and A.C. Kraay (1998) Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80: 549-560.

Elhorst J.P., D.J. Lacombe and G. Piras (2012) On model specifications and parameter space definitions in higher order spatial econometric models. *Regional Science and Urban Economics*, 42: 211-220.

Helpman E, M. Melitz and Y. Rubinstein (2008) Estimating trade flows: trading partners and trading values. *Quarterly Journal of Economics*, 132: 441-487.

Hicks J.R. (1932) *The Theory of Wages*, London, Macmillan.

LeSage J.P and M.M. Fischer (2010) Spatial econometric methods for modeling Origin – Destination flows. *Handbook of Applied Spatial Analysis: SoftwareTools, Methods and Applications*, M.M. Fischer and A Getis (eds), Springer-Verlag, Berlin Heidelberg.

LeSage J. and R.K. Pace (2008) Spatial econometric modeling of origin-destination flows. *Journal of Regional Science*, 48: 941-967.

McKinnish T. (2007) Welfare-induced migration at state borders: new evidence. *Journal of Public Economics*, 91: 437-450.

Péridy N. (2006) The European Union and its new neighbors: an estimation of migration potentials. *Economics Bulletin*, 6: 1-11.

Pöyhönen P. (1963) A tentative model for the volume of trade between countries. *Wetwirtschaftliches Archiv (Review of World Economics)*, 90: 92-100.

Razin A., E. Sadka and B. Suwankiri (2011) *Migration and the Welfare State*, Cambridge MA, MIT Press.

Sjaastad L.A. (1962) The costs and returns of human migration. *Journal of Political Economy*, 70: 80-93.

Smith A. (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*, K. Sutherland (ed), Oxford University Press, 1993.

Willcocks G (2010), Conditional Variances in UK Regional House Prices, *Spatial Economic Analysis*, 5(3), 339-354

Yoshida C. and A.D. Woodland (2005) *The Economics of Illegal Immigration*, New York, Palgrave Macmillan.