# Spatial Dependence in the Econometrics of Gravity Modeling

**Michael Beenstock**
**Daniel Felsenstein**

**Hebrew University of Jerusalem**

11th July 2013

**Introduction: 150 Years of Gravity**

Gravity modeling emerged over a century ago as an attempt to harness Newtonian physics in the explanation of socio-economic processes. Ravenstein's Law of Migration (1885) and Reilly's Law of Retail Gravitation (1929) are but two examples of the mechanistic straightjacket of early social physics. With the expansion of applications to spatial consumer choice, commuting patterns and housing choice, a more behavioral gravity model emerged. This embraced the principles of minimum effort (Zipf 1949) intervening opportunities (Stouffer 1940) and demographic potential (Stewart 1948) . Over time, the use of gravity models in spatial analysis veered away from social physics and contemporay spatial gravity modeling is now part of a toolkit of spatial interaction techniques that run from entropy maximization (Wilson 1971) through to neural network modeling (Fischer, Reismann and Hlavackova-Schindler 2003).

A major juncture in the development of gravity modeling developed 50 years ago in the field of bilateral trade flows with the pioneering work of Tinbergen (1962) and Pöjhönen (1963). In its basic form, the gravity model hypothesizes that bilateral transactions between origins and destinations vary inversely with the distance between them, as well as with pull factors in destinations and push factors in origins. Although gravity modeling was initially applied to international trade, it was subsequently extended to the study of international capital flows and international migration. Gravity modeling has also been applied intranationally, e.g. in the study of internal migration. Indeed, the gravity model has served as a methodological work-horse in numerous empirical studies involving origins and destinations. Most probably the number of published papers using gravity modeling runs into the thousands.

Although the basic gravity assumption, the strength of attraction between origins and destinations varies inversely with distance between them, makes intuitive sense, it was not until the late 1970s that the theoretical underpinnings of gravity in international trade were formulated (Andersen 1979). Subsequently trade theorists have disputed whether gravity is consistent with the old theory of international trade based on Hecksher-Ohlin or the new theory of international trade based on imperfect competition (Bergstrand 1985, 1989, Deardorff 1998, Everett and Keller 2002).

It took another 20 years for the theoretical underpinnings of gravity in international trade and migration to be formulated in terms of Multilateral Resistance (Anderson and van Wincoop 2003, 2004 and Feenstra 2004), according to which traders or migrants face a discrete choice problem in choosing to trade with or emigrate to alternative destinations. The common denominator to these theories is that national markets in the case of trade, and domiciles in the case of migration are imperfect substitutes and that trade and migration involve frictions. In this chapter, however, our concern does not lie with gravity theory but with its econometric aspects. Surprisingly, the latter have attracted little attention, except until recently.

In gravity models the dependent variable is a bilateral flow from an origin to a destination. If there are N locations or nodes there must be N(N-1) bilateral observations. The standard econometric assumption made in innumerable studies has been that these observations are independent, which enables the use of ordinary least squares (OLS) to estimate the parameters of the gravity model. Denoting the residuals from the gravity model by $u_{od}$ (where o labels origins and d labels destinations), OLS assumes that $u_{od}$ is independent of $u_{do}$. For example, Italian exports to Israel are independent of Italian imports from Israel. This assumption may be contravened for a variety of reasons[1]. OLS also assumes the $u_{od}$ is independent of $u_{os}$ where s is another destination. If, for example s refers to Greece and the economies of Israel and Greece are related directly through international trade or indirectly through third countries, Italian exports to Israel may not be independent of Italian exports to Greece. OLS also assumes that Israel’s exports to Italy are independent of Israel’s exports to third countries such as France. In short, the assumption that the residuals are independent may be incorrect.

If the gravity residuals are dependent OLS estimates of the gravity model parameters are inefficient but consistent. Since this issue has been ignored in the literature, there may be many results that are incorrectly reported as being statistically significant. In principle robust standard errors may be calculated which take account of the dependence between the gravity residuals. Driscoll and Kraay (1998) have suggested such a procedure for spatially correlated residuals. More generally, the solution to this problem would be seemingly-unrelated regression (SUR) in which the estimated standard

[1] For example, Israel swimsuit exports use fabrics made in Italy.

errors are calculated under the assumption that the gravity residuals are dependent. However, SUR is only feasible in the case of panel data.

The issue of dependence runs deeper than this; it does not merely concern the gravity residuals, but the specification of the gravity model itself. Since trade is essentially multilateral, a bilateral specification is miss-specified. The trade flows between Italy and Israel do not only depend on push and pull factors in these countries, but also on these factors in third countries. For example, an increase in GDP in France might affect bilateral trade between Italy and Israel. In principle, the gravity model should specify all N – 2 push-pull factors. Since in practice gravity models do not specify third country effects, numerous studies may have omitted variables that are empirically relevant. If these omitted variables are correlated with the variables in the gravity model, the parameter estimates of the gravity model will be biased and inconsistent. This criticism applies to hundreds of studies that have been published during the last 50 years.

Econometric theory for gravity modeling only began to receive attention in the last few years. LeSage and Pace (2008) were the first to draw attention to the problem. Although the problem is essentially multilateral, LeSage and Pace assume that the data are spatially dependent. This simplification enables them to draw upon spatial econometrics by specifying spatially lagged dependent variables in the gravity model, and by allowing the gravity residuals to be spatially autocorrelated. They specify separate spatial connectivity matrices for origins and destinations. If multilateralism happens to be spatial this solution is fine. However, it might not be. In the case of trade, for example, Israel's high-tech exports to Italy may be multilaterally related to Israel's competitors in the US and Finland, which are remote, rather than to Israel's immediate neighbors in the Middle East. Behrens, Ertur and Koch (2012) have adapted LeSage and Pace's ideas to multilateral resistance theory by giving spatial connectivity matrices a multilateral rather than a spatial interpretation[2]. Recently this attention to spatial dependence has been extended to the case where 'latent' spatial effects are estimated for both the origin and destination (LeSage and Llano 2013). This involves the estimation of  a Bayesian

[2] Behrens et al assume that because income and the number of product varieties vary directly with scale, larger economies are more likely to trade with each other than smaller ones. According, spatial weights are defined in terms of the relative size of regions as reflected in population shares. Their identification strategy assumes that internal migration is independent of trade.

hierachical model that uses the SAR structure as a spatial prior to structure the regional effects parameters.

It should be clear by now that econometric theory for gravity has lagged substantially behind the economic theory of gravity. Indeed, the econometric theory for gravity is in its infancy. Economists have been preoccupied instead with other econometric problems that arise in gravity models, and especially how to deal with the fact that many bilateral flows are zero, and their implications for testing hypotheses about extensive and intensive margins. Helpman, Melitz and Rubinstein (2008) specify a probit selection model for zero trade flows, and Burger, van Oort and Linders (2009) apply a zero-inflation methodology. In our opinion this a second order problem; the main methodological problem stems from the fact that gravity is essentially multilateral rather than bilateral. Because this problem has been ignored a cloud of doubt hovers over countless empirical studies, some of them influential, based on OLS estimates of bilateral gravity.

Dependence between gravity residuals affects the efficiency but not consistency of OLS estimates of gravity parameters. Matters are different regarding the nonlinear maximum likelihood estimators used to handle zero bilateral flows. Dependence between residuals in probit and zero-inflating estimators induces inconsistency. Having solved one problem, Helpman et al and others might have created another. It is difficult to know whether OLS that ignores zero bilateral flows is inferior to ML which does not ignore zeros, but which ignores dependence between gravity residuals.

Griffith and Fischer (2013) suggest treating spatial dependence by using spatial filtering. This involves screening the sample origin-destination data for spatial association by transferring SAC effects from residuals to the mean or intercept. This creates “spatially adjusted” data suitable for Poisson estimation (ie count based). The Poisson regression interprets the flows as dependent on the origin and destination-specific effects coefficients. Spatial filtering implicitly assumes that spatiality is a nuisance parameter that may be “concentrated out” to estimate the parameters of interest. Like seasonality in time series data, spatiality may not be independent of the parameters of interest. Indeed, since trade and migration are inherently multilateral, spatial filtering may treat parameters of interest as nuisances.

In this chapter we make the following methodological contributions to the econometrics of gravity modeling. First, we consider the case in which origins and destinations are not mutual, i.e. countries or spatial units are either origins or destinations, but not both, so bilateral flows are one way. Second, we propose a lagrange multiplier test for spatial autocorrelation among origins and among destinations. Third, we also propose a test for spatial autoregressive conditional heteroscedasticity (SpARCH) between origins and destinations. SpARCH exists when error variances are spatially autocorrelated; it is the spatial counterpart to ARCH in time series, and it is the counterpart to spatial autocorrelation for variances[3]. More generally, whereas spatial econometric analysis has been almost exclusively concerned with spatial dependence between means, as e.g. in the spatial lag model, we draw attention to potential spatial dependence between second moments as well as between first moments. We illustrate these concepts empirically with an application to migration from European Neighborhood Countries (EN) to members of the European Union (EU). Since migration from EN to EU is one-way, EN countries are origins and EU countries are destinations.

A limitation is that multilateralism is assumed here to be spatial. This means, for example, that when Egyptians chose to emigrate to France their decision is not independent of local alternative destinations to France, such as Germany. However, it is independent of distant alternatives, such as the United States. It also means that when Libyans chose to emigrate to France their decision is not independent of their Egyptian neighbors' decisions to emigrate to France. However, their decisions are independent of emigration decisions in origins remote from Libya, such as Ukraine. The implicit assumption in spatial multilateralism is that, everything else given, destinations are closer substitutes the nearer they are, and that shocks are likely to be more correlated among origins the closer they are

This implicit assumption is no doubt too restrictive because multilateralism is not merely spatial. Quebec may be a closer substitute to France for francophone Algerians than Germany regardless of distance. Also, the emigration decisions of Israelis

[3] SpARCH is not to be confused with the spatial GARCH model in Willcocks (2010) in which the variance in location i at time t depends on the variance of location j at time t-1. Nor should it be confused with the SEARCH model of Caporin and Paruolo (2005) in which the residuals are spatially autocorrelated in a regular ARCH model, i.e. $u_{it} = Wu_{it} + e_{it}$ where e is an ARCH process.

and Egyptians are unlikely to be correlated just because they happen to be in the Middle East. If decision making in migration and trade is hierarchical or nested, then spatial effects are likely to be important. However, if it is direct and unmediated, then we can assume relations are multilateral.

**2. Origins and Destinations**

Let $y_{od}$ denote the bilateral flow between origin o and destination d. There are $N_o$ origins, $N_d$ destinations and $N = N_dN_o$ one-way bilateral flows. Let $y_o$ denote an $N_d$-length vector of bilateral flows from origin o to all destinations. These vectors are stacked, as in panel data, to form an N-length vector of bilateral flows y:

$$y' = (y'_1 \; y'_{2.} .... y'_{N_o})$$

The first $N_d$ elements of y refer to the flows of origin 1 to all destinations. $W_o$ is a square $N_o$-matrix with zeros along the leading diagonal of spatial weights in the origins:

$$W_o = \begin{bmatrix} 0 & w_{12} & .. & w_{1N_0} \\ w_{21} & 0 & .. & w_{2N_0} \\ .. & .. & .. & .. \\ w_{N_o1} & w_{N_o2} & .. & 0 \end{bmatrix}$$

and $W_d$ is a square $N_d$-matrix of spatial weights in the destinations:

$$W_d = \begin{bmatrix} 0 & w_{12} & .. & w_{1N_d} \\ w_{21} & 0 & .. & w_{2N_d} \\ .. & .. & .. & .. \\ w_{N_d1} & w_{N_d2} & .. & 0 \end{bmatrix}$$

Define $D = I_{N_o} \otimes W_d$ and $\Omega = W_o \otimes I_{N_d}$, which are NxN matrices. D is block diagonal with $W_d$ along the leading diagonal and zeros elsewhere. Ω has zeros along the leading diagonal and $w_{od}I_{Nd}$ elsewhere. The vector of spatial lags in origins and destinations may be defined as:

$$\tilde{y}_o = \Omega y$$
$$\tilde{y}_d = Dy$$
$$\tilde{y} = \tilde{y}_o + \tilde{y}_d$$

For example, $y_{od}$ is the flow from origin o to destination d. Let o be Egypt and d be France. Flows from Egypt to France might be related to flows from Libya (Egypt’s neighbor among origins) to France. This spatial lag component is included in $\tilde{y}_o$ because Libya and Egypt are origins with common destinations. Flows from Egypt to France might be related to flows from Egypt to Germany. This spatial lag component is included in $\tilde{y}_d$ because France and Germany are destinations with common origins.

The generalized spatial lag model with origins and destinations (GSOD) is:

$$y = X_o^*\alpha + X_d^*\beta + \gamma C + \tilde{X}_o^*\delta + \tilde{X}_d^*\pi + \rho_o\tilde{y}_o + \rho_d\tilde{y}_d + e \qquad (1)$$

$$X_o^* = \mathrm{I}_{N_d} \otimes X_o$$

$$X_d^* = \mathrm{I}_{N_o} \otimes X_d$$

$$\tilde{X}_o^* = \mathrm{I}_{N_d} \otimes W_o X_o$$

$$\tilde{X}_d^* = \mathrm{I}_{N_o} \otimes W_d X_d$$

where C is a vector of distances between origins and destinations with elements $c_{od}$, $X_o$ is an $N_o x K_o$ matrix of push factors in the origins, and $X_d$ is an $N_d x K_d$ matrix of pull factors in the destinations. Their spatial counterparts are $\tilde{X}_o = W_o X_o$ and $\tilde{X}_d = W_d X_d$, α and δ are $K_o$-length vectors of parameters, β and π are $K_d$-length vectors of parameters, and $\rho_o$ and $\rho_d$ are spatial lag coefficients in origins and destinations. Finally, e is an N-vector of residuals. Equation (1) states that flows e.g. from Egypt to France depend on push factors in Egypt through α, push factors in Libya through δ, pull factors in France through β, and pull factors in Germany through π. They also depend on flows from Libya to France via $\rho_o$, and from France to Germany via $\rho_d$.

Let $x_{di}$ denote a pull factor in destination i. According to GSOD the partial derivative of $x_{di}$ on bilateral flows is:

$$\frac{dy}{dx_{di}} = \left(\mathrm{I}_N - \rho_o\Omega - \rho_d D\right)^{-1}\left(\beta + \pi \mathrm{I}_{N_o} \otimes W_d\right) \qquad (2)$$

$$(\mathrm{I}_N - \rho_o\Omega - \rho_d D)^{-1} = \mathrm{I}_N + \rho_o\Omega + \rho_d D + \rho_o^2\Omega^2 + \rho_d^2 D^2 + \rho_0\rho_d\Omega D + .. \quad (2a)$$

An increase in $x_{di}$ directly pulls flows from all origins to i. For example, an increase in GDP in France induces flows into France from all origins via β. The increased flow from Libya to France induces an additional increase from Egypt to France via $\rho_o$. The increased flow from Egypt to France affects flows from Egypt to Germany via $\rho_d$. The increase in French GDP has an independent affect of flows from Egypt (and other origins) to Germany via $\pi$. Since each unit is its neighbor's neighbor among origins and destinations these effects propagate further. These effects are expressed in equation (2a) in which $\Omega^2 = W_o^2 \otimes \mathrm{I}_{N_d}$, $D^2 = \mathrm{I}_{N_o} \otimes W_d^2$ where $W_o^2$ and $W_d^2$ refer to second order neighbors in origins and destinations, while the final term $\Omega D = W_o \mathrm{I}_{N_o} \otimes W_d \mathrm{I}_{N_d}$ refers to interactions between $W_o$ and $W_d$. Because $\rho_o$ and $\rho_d$ are less than one in absolute value (otherwise y would be spatially nonstationary, Beenstock, Feldman and Felsenstein 2012), equation (2a) has a finite inversion since powers of $\rho_o$ and $\rho_d$ and their product tend to zero. In short, GSOD specifies a rich range of spatial dynamics of the autoregressive and moving average varieties through $\rho_o$ and $\rho_d$, and $\delta$ and $\pi$ respectively.

## 3. The Econometrics of Spatial Gravity Modeling

### *3.1 Double Spatial Lagged Dependent Variables*

Since GSOD involves a double spatial lag, estimation is not straightforward because the likelihood function involves the determinant $|\mathrm{I}_N - \rho_o \Omega - \rho_d D|$. If $\rho_o = \rho_d$ matters are simplified and the determinant reverts to its standard form involving a single spatial lag, in which case standard estimators available in Matlab etc may be used. The likelihood has to be maximized with respect to $\rho_o$ and $\rho_d$ as well as other GSOD parameters. We use the double spatial lag estimator developed by Elhorst et al (2012) to estimate the parameters of GSOD.

### *3.2 Spatial Autocorrelation*

The GSOD residuals (e) are assumed to be iid random variables that are asymptotically normal. Spatial autocorrelation in GSOD residuals may arise either because the residuals are spatially correlated among origins, or because they are spatially correlated among destinations. For example, spatial autocorrelation among origins arises when the residuals for Egyptian flows to France and other destinations are correlated with Libya's residuals with respect to France as well as other destinations. Spatial autocorrelation among destinations arises when the residuals for Egyptian flows to France are correlated with Egypt's residuals with respect to Germany and other destinations.

We suggest the following auxiliary regression to test for both types of spatial autocorrelation:

$$\hat{e} = Z\psi + \lambda_o \tilde{e}_o + \lambda_d \tilde{e}_d + \varepsilon \qquad (3)$$
$$\tilde{e}_o = \Omega \hat{e}$$
$$\tilde{e}_d = D\hat{e}$$

where Z is a matrix of the covariates used to estimate GSOD, ê are the GSOD residuals estimated by ML, and ε is iid. The absence of spatial autocorrelation means that $\lambda_o$ and $\lambda_d$ are zero, in which case ψ must be zero. The lagrange multiplier statistic is LM = $NR^2$ where $R^2$ is for equation (3). It has a chi-square distribution with 2 degrees of freedom for the two independent restrictions regarding $\lambda_o$ and $\lambda_d$.

If the GSOD residuals happen to be spatially autocorrelated, this may indicate that the GSOD model is spatially misspecified, or it may suggest that it is correctly specified but the residuals just happen to be spatially autocorrelated. A straight-forward common factor test may be used to distinguish between these alternatives. In the former case, if $\lambda_o$ only is statistically significant the spatial misspecification arises among the origins, and if $\lambda_d$ only is statistically significant the spatial misspecification arises among the destinations.

*3.3 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. Vectorizing equation (1) we rewrite it as:

$$y = Q\omega + e \qquad (4$$
$$e = (\lambda_o \Omega + \lambda_d D)e + \varepsilon \qquad (5)$$

where Q refers to the regressors in equation (1) and ω their coefficients[4]. The solution to equation (5) is:

$$e = A\varepsilon \qquad (6)$$
$$A = (I_N - \lambda_o \Omega - \lambda_d D)^{-1}$$

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

[4] Q excludes spatial lagged dependent variables, otherwise OLS would not be appropriate.

$$\Sigma_{\omega} = (Q'Q)^{-1}(Q'\Theta Q)(Q'Q)^{-1} \qquad (7)$$
$$\Theta = A\Sigma_{\varepsilon}A'$$

If ε is homoscedastic $\Theta = \sigma_{\varepsilon}^{2}AA`$. To implement equation (7) estimates of A and $\Sigma_{\varepsilon}$ based on estimates of $\lambda_o$, $\lambda_d$ and ε obtained from equation (3) are substituted into equation (7). If ε is heteroskedastic $\Theta = A\Xi A`$ where Ξ is a diagonal matrix with diagonal elements $\hat{e}_{od}^{2}$ .

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 $\lambda_o$ and $\lambda_d$.

### *3.4 Spatial Autoregressive Conditional Heteroskedasticity*

Another type of potential dependence concerns variances. We suggest that the spatial counterpart to the ARCH (autoregressive conditional heteroskedasticity) that arises in time series may be specified as:

$$\hat{e}^{2} = \omega + \omega_{o}\tilde{e}_{o}^{2} + \omega_{d}\tilde{e}_{d}^{2} \qquad (8)$$

The spatial ARCH (SpARCH) parameters are $\omega_o$ and $\omega_d$, which might differ between origins and destinations. Equation (8) assumes that volatility may be transmitted spatially, and that the conditional variance of $e_{od}$ depends on volatility in the vicinity of o among origins, and in the vicinity of d among destinations. These variances are therefore conditionally heteroskedastic. By contrast, the unconditional variance is:

$$\sigma_{e}^{2} = \omega\left[\mathrm{I}_{N} - \omega_{o}\Omega - \omega_{d}D\right]^{-1}$$

Since this does not depend on o or d the unconditional variance is the same for all variance. The LM test for SpARCH involves using the estimated GSOD residuals to estimate equation (8). The test statistic is $NR^2$ and has a chi-squared distribution with 2 degrees of freedom.

Whereas unconditional homoskedasticity is one of the classical assumptions required for OLS, conditional heteroskedasticity does not violate these assumptions. Therefore, evidence of SpARCH does not matter in OLS contexts. With nonlinear estimators matters are different; ARCH interferes with consistency. Since the spatial lag parameters in GSOD obtained by ML are nonlinear, SpARCH induces inconsistency in the estimates of the GSOD parameters.

## 4. The European Neighborhood

The European Neighborhood (EN) is a geopolitical concept (see map) as defined by EU foreign policy in general and the European Neighborhood Policy (ENP) in particular[5]. It includes countries that are not candidates for EU membership, hence Turkey is not included in the EN. EN includes all countries in North Africa with coasts on the Mediterranean. It includes countries in the Middle East (Israel, Jordan, Lebanon and Syria), countries in South Caucasia (Georgia, Armenia, Azerbaijan), and countries in the former USSR (Ukraine, Belorussia and Moldava), making 16 countries in all. The EU regards EN countries as their political and economic hinterland. These EN countries serve as origins in the present study.

Dealing with migration flows from the ENP countries is high on the EU policy agenda. The EU shares a 5000+km border with the ENP countries to the east and a similar length (albeit maritime) border with the ENP countries to the south. EU policy relating to migration from the ENP countries has been articulated in various agreements such as the Amsterdam Treaty and the Tampere , Hague and Stockholm Programs. Migration policy with respect to ENP countries is part of an EU attempt to regulate border security in three areas: illegal (or irregular) migration, combating trafficking and smuggling of human beings and cross-border management practice. Ostensibly, regulated migration policy is perceived as benefitting both origin and destination countries. For the ENP countries, migration is a solution for the lack of local employment opportunities. For the EU countries, it provides a solution to the demographic imbalance and ageing population trends in the core countries over the short run. Potentially, migration policy could be conceived as diverting human disaster in the ENPs and promoting growth and prosperity in the EU.

The EU currently has 28 members, including countries such as Latvia, Romania and Croatia that have joined recently. In principle these countries serve as the destinations. However, we restrict the EU destinations to the 15 members prior to the recent enlargement for two reasons. First, the study period refers to immigration during

[5] ENP involves concessions to EN countries regarding trade, investment and migration. It also obliges neighboring countries to adapt local legislation to EU norms thereby extending integration without formal enlargement (Harpaz, 2013).

the first decade of the 21st century. Since countries such as Romania and Bulgaria were not members in 2000 they are omitted from the study. Secondly, it turns out that there were no immigrants from EN in the 10 omitted EU members. Dropping these countries conveniently means that we may ignore the problem of treating zero bilateral flows. Therefore, $N_o$ is 16 and $N_d$ is 15.

We use the Global Bilateral Migration Database (GBMD, World Bank) which provides estimates of the number of foreign born by all origins of the world in all destinations[6]. Table 1 presents these data in 2000 for our 16 origins in the 15 destinations. Notice that with some exceptions in Portugal these population stocks are non-zero. The number of foreign born is not necessarily equal to the number of immigrants because they include children[7]. GBMD is decennial starting from 1960. Since GBMD refers to population stocks, we define immigration flows from origins to destinations by the stock in 2010 minus the stock in 2000. GBMD in principle covers people who returned to their country of origin by 2010 or migrated to third countries. However, foreign born who died between 2000 and 2010 would be registered as a decrease in the number of foreign-born. Therefore, our definition of immigration flows is an under-estimate because GBMD does not identify the deceased.This partly explains why the estimated flows of immigrants (Table 2) are occasionally negative. The other reason might be data errors in GBMD. Table 2 expresses the changes in immigrant (foreign born) stocks during the first decade of the 21st century as a percentage of the stock in 2000. Some of these estimated rates of immigration are very large especially in destinations where the initial stock was small (e.g. Portugal)

.

## 5. Immigration Theory

This paper tests the welfare-motivation pull factor hypothesis of migration.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

[6] See Özden et al (2011) for methodological details how GBMD was constructed.
[7] Data for Israel in GBMD differ to data published by Israel's Central Bureau of Statistics. We have been unable to obtain an explanation for this large discrepancy from the World Bank.

internal migration in England immigration as unjust and economically harmful[8]. 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[9]. 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 deterred by the higher taxation required to finance this generosity. In all of these studies it is assumed that bilateral migration flows are independent.

*5.1 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.

Let $Y_{odt}$ denote the stock of immigrants from origin o in destination 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:

---

[8] Smith (1976) argued that the Law of Settlements, enacted to enforce poor law benefits provided 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.

[9] 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.

$$y_{odt} = \phi(Y^*_{odt} - Y_{odt}) + \varphi\Delta Y^*_{odt+1} \quad (9)$$

where ϕ and φ are stock adjustment coefficients. Let P denote an $N_d x K_d$ matrix of pull factors in the destinations, and U denote an $N_o x K_o$ matrix of push factors in the origins. In principle, immigrants from o may choose between all destinations. Therefore:

$$Y^*_t = P_t\zeta + U_t\xi \quad (10)$$

Where Y* is an N-vector, ζ and ξ are vectors of parameters of length $N_d$ and $N_o$ respectively. Equation (10) states that the equilibrium number of immigrants from o in d at time t depends via ζ on the pull factors in the destinations, and via ξ on the push factors in the origins. Substituting equation (10) into equation (9) gives:

$$y_t = \phi(P_t\zeta + U_t\xi) + \varphi(\Delta P_{t+1}\zeta + \Delta U_{t+1}\xi) - \phi Y_{odt} \quad (11)$$

Therefore in equation (1) $X_o = P_t + \Delta P_{t+1}$ and $X_d = U_t + \Delta U_{t+1}$. Equation (11) 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 (11) they compare pull factors in France with pull factors in other EU countries.

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 (11).

*5.2 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 wage 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 or irregular immigrants is more complicated. Procedures for dealing with political refugees vary by country; they may be more or less lenient. If country 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[10].

We have collected data on the rights of legal and irregular immigrants, as well as on the way countries treat irregular 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 (MIPEX) of the treatment of legal immigrants in EU destinations in terms of the assistance they get to integrate economically, socially and politically.

## 6. 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 2. The origin variables ($X_o$) include GDP per head in 2000 and its rate of growth during 2000 – 2010. The destination variables ($X_d$) 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

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

between origins and destinations, common official languages, and immigrant stocks in 2000.

Most of these variables did not turn out to be 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 ENP 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.

In sum, these results do not point to welfare generosity in EU destinations as influencing migration from the ENP's. Nor do enforcement measures against irregular immigrants seem to deter migration from the ENP's. Although the evidence is not strong enough to support substantive policy prescriptions, it does imply that reduced economic growth in EU and cuts in welfare are unlikely to reduce the flow of immigration from ENP countries. On the other hand, the influence of neighboring countries seems to be of more importance. Immigration to the EU is positively influenced by immigration to its neighbors and vice-versa. Emigration from an ENP country to the EU is positively influenced by emigration from its neighbors and vice-versa. These powerful spatial spillovers suggest that piecemeal immigration measures are unlikely to succeed and that EU actions that encourage immigration from specific ENP countries may induce immigration from these countries neighbors.

## 7. Conclusions

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 20$^{th}$ 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. Social welfare policy and policy towards illegal migrants in EU destinations do not seem to impact the flow of migrants from the ENP countries. The paper also addresses the extent to which destination choices within the EU are complements and substitutes. This has policy ramifications with respect to the spillover of migration pressure points within the EU. A parochial policy which, for example, restricts migration in one country might deflect immigration to its neighbors. Also a policy which encourages immigration in one country might induce immigration to its neighbors. Thus immigration policy would need to be designed globally rather than parochially.

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.

# Immigrants 2000

| Origin \ Destination | 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 |

# Immigration rates (%)2000-2010

| Destination / Origin | Austria | Belgium | Denmark | Finland | France | Germany | Greece | Ireland | Italy | Luxem- -bourg | Netherlands | Portugal | Spain | Sweden | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Algeria | 34.4 | 169.2 | 27.4 | 68.6 | -13.6 | 3.9 | 41.6 | 125.9 | 85.9 | 16.7 | -1.0 | | 172.2% | 33.6 | -61.5 |
| Armenia | -9.9 | 491.3 | 30.8 | 66.3 | 389.1 | -28.5 | 17.9 | 159.6 | 97.9 | 16.7 | 658.7 | 310.5 | 395.1% | 98.2 | 5160.0% |
| Azerbaijan | 40.0 | 753.8 | 50.4 | 68.3 | -8.4 | 1032.1 | 52.9 | 160.5 | 156.6 | 25.0 | 566.9 | 700.0 | 295.1% | 117.3 | 37450.0 |
| Belarus | 48.5 | 1102.2 | 134.3 | 66.2 | 36.3 | 664.1 | 61.0 | 88.2 | 230.1 | 21.4 | 625.4 | 3620.0 | 474.1% | 120.8 | 3260.9 |
| Egypt | 79.9 | 258.6 | 37.3 | 80.2 | 453.8 | 47.0 | 28.9 | 134.8 | 108.1 | 20.6 | 20.5 | | 156.6% | 36.9 | 4.2 |
| Georgia | 98.5 | 63.8 | 43.6 | 57.4 | -92.5 | -75.8 | 90.3 | 144.0 | 313.8 | 16.7 | 732.7 | 89.5 | 698.1% | 121.3 | 800.0 |
| Israel | 27.8 | 126.6 | 40.8 | 80.5 | 77.3 | 50.6 | 124.2 | 128.4 | 18.4 | 21.6 | 20.7 | | 225.9% | 45.7 | 75.1 |
| Jordan | 32.0 | 115.6 | 38.8 | 80.5 | 51.0 | 42.4 | 49.7 | 144.5 | 24.7 | 16.7 | 5.4 | | 96.9% | 50.6 | 548.7 |
| Lebanon | 175.6 | 332.2 | 28.3 | 79.5 | 312.0 | 19.3 | 206.1 | 148.3 | 143.7 | 20.7 | 9.6 | | 110.9% | 23.3 | 39.2 |
| Libya | 17.9 | 549.2 | 38.9 | 82.4 | 268.3 | 437.8 | 55.9 | 138.7 | -41.7 | 20.0 | 26.8 | | 293.8% | 49.5 | 8802.9 |
| Morocco | 41.5 | 55.6 | 34.4 | 59.5 | 220.4 | 28.2 | 36.1 | 95.0 | 66.1 | 19.7 | 10.6 | 74.5 | 207.5% | 40.5 | -40.2 |
| Moldova | 45.5 | 167.4 | 74.3 | 66.2 | -72.1 | 26.9 | 34.4 | 248.1 | 1235.1 | 20.0 | 590.9 | 45.5 | 857.5% | 173.2 | 238.3 |
| Russia | 77.4 | 2794.5 | 91.8 | 66.6 | -80.2 | -69.4 | 125.4 | 150.0 | 88.3 | 18.9 | -74.9 | 124.8 | 442.2% | 58.1 | 121.3 |
| Syria | 162.8 | 323.9 | 71.0 | 80.3 | 192.3 | 54.5 | 99.1 | 105.9 | 34.9 | 18.2 | 18.4 | | 100.5% | 38.5 | -2.2 |
| Tunisia | 60.7 | 195.8 | 34.2 | 76.4 | -2.8 | 46.7 | 58.2 | 99.2 | 60.5 | 19.4 | 11.4 | | 170.7% | 33.8 | -59.1 |
| Ukraine | 68.7 | 265.4 | 486.4 | 66.6 | 29.6 | 248.2 | 89.8 | 221.1 | 1154.6 | 20.6 | 610.2% | 56.8 | 377.4% | 76.8 | 3090.2 |

**Table 3 Estimates of the Migration Model: 2000-2010**

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

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<br>(0.23) | -0.4768<br>(-2.06) | -0.9941<br>(-9.16) |
| Destination | -0.0511<br>(-0.63) | -0.0840<br>(-0.37) | -0.9725<br>(-8.90) |
| LM | 2.6015 | 24.209 | 81.399 |
| SpARCH | | | |
| Origin | 0.6596<br>(0.59) | 0.9152<br>(4.18) | 0.5922<br>(4.33) |
| Destination | 0.0167<br>(0.25) | 0.2350<br>(2.44) | 0.5961<br>(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 $\chi^2$ (df = 2) = 5.991. t-statistics for SAC and SpARCH coefficients reported in parentheses.

**Appendix: 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 | MIPEX – Migrant Integration Policy Index | http://www.mipex.eu/sites/default/files/down |

| | | | | |
|---|---|---|---|---|
| | | reunion of immigrants | | loads/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 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 |
| Housing | Index | Index based on the level of punishment for renting shelter to | FRA (European Union Agency for Fundamental Rights) - | http://research.icmpd.org/fileadmin/Research-h- |

| | | | | |
|---|---|---|---|---|
| | | migrants in an irregular situation | Fundamental rights of migrants in an irregular situation in the European Union | 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%20on%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 | EMN (European | http://emn.intrasoft- |

| | | | | |
|---|---|---|---|---|
| | | foreign nationals removed vs. the migrant stock in the destination country | Migration Network) - Annual Report on Migration and International Protection Statistics 2003-2009 | intl.com/Downloads/prepareShowFiles.do?entryTitle=2%2E%20Annual%20Reports%20on%20Migration%20and%20International%20Protection%20Statistics |
| | | | | |
| | | | | |

**References**

Anderson J.E. (1979) A theoretical foundation for the gravity equation. *American Economic Review*, 69: 105-116.

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

Beenstock M., D. Feldman and D. Felsenstein (2012) Testing for unit roots and cointegration in spatial cross-section data. *Spatial Economic Analysis*, 7: 203-222.

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.

Bergstrand J.H. (1985) The gravity equation in international trade: some microfoundations and empirical evidence. *Review of Economics and Statistics*, 67: 474-481.

Bergstrand J.H. (1989) The generalized gravity equation, monopolistic competition, and the factor proportions theory of international trade. *Review of Economics and Statistics*, 71: 143-153.

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

Caporin M. and P. Paruolo (2005) Spatial effects in multivariate ARCH, discussion paper, Universita dell' Insurbia, Varese.

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

Deardorff A.V. (1998) Determinants of bilateral trade: does gravity work in a neoclassical world? In J,A. Frankel (ed) *Regionalization of the World Economy*, University of Chicago Press.

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.

Evenett S. and W. Keller (2002) On theories explaining the gravity equation. *Journal of Political Economy*, 110: 281-316.

Feenstra R.C. (2004) *Advanced International Trade: Theory and Evidence*, Princeton University Press.

Feenstra R.C., J.R. Markusen and A.K. Rose (2001) Using the gravity equation to differentiate among alternative theories of trade. *Canadian Journal of Economics*, 34: 431-447.

Fischer, M.M., Reismann, M. and Hlavackova-Schindler, K. (2003). Neural network modelling of constrained spatial interaction flows: Design, estimation and performance issues. *Journal of Regional Science*, Vol. 43(1), 35-61

Griffith D.A. and Fischer M.M. (2013) Constrained Variants of the Gravity Modle and Spatial dependence: Modle Specification and Estimation Issues, *Journal of Geographic Systems*, 15, 291-317.

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 Llano C (2013) A Spatial Interaction Model with Spatially Structured Origin and Destination Effects, *Journal of Geographic Systems*, 15, 265-289.

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.

Özden C., C.R. Parsons, M. Schiff and T. Walmsley (2011) Where on earth is everybody? The evolution of global bilateral migration 1960 – 2000. *World Bank Economic Review*, 25: 12-56.

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.

Ravenstein (1885) The Laws of Migration, *Journal of the Royal Statistical Society* , 48, 163-235.

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

Reilly WJ (1929) *Methods for the Study of Retail Relationships*, University of Texas Bulletin, 2944, University of Texas Austin.

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.

Stewart JQ (1948) Demographic Gravitation: Evidence and Application, *Sociometry*, 11, 31-58

Stouffer A (1960) Intervening Opportunities and Competing Migrants, *Journal of Regional Science* 1, 1-20.

Tinbergen J. (1962) An analysis of world trade flows. In J. Tinbergen (ed) *Shaping the World Economy*, New York, Twentieth Century Fund.

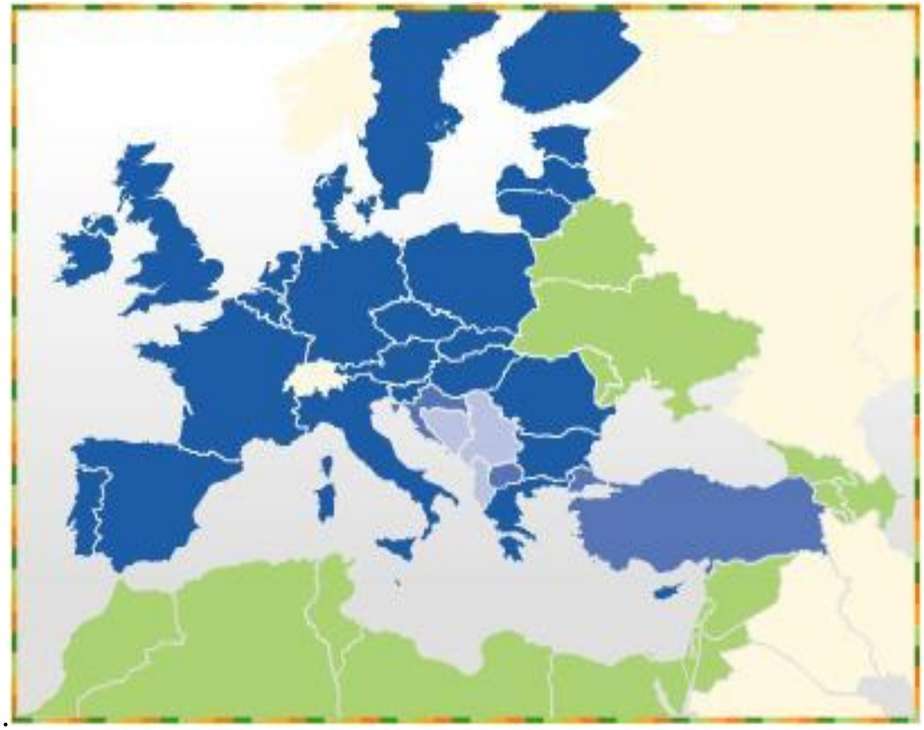
Willcocks G (2010), Conditional Variances in UK Regional House Prices, *Spatial Economic Analysis*, 5(3), 339-354

Wilson AG (1971) A Family of Spatial Inteaxction Models and Associagted Developments, *Environment and Planning A*, 3, 1-32

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

Zipf GK (1949) *Human Behavior and the Principle of Least Effort*, Addison-Wesley, Reading, MA

Figure 1; The European Neighborhood; blue (EU), green (ENP)



.