Do Migrants Follow Market Potentials? An Estimation of a New Economic Geography Model

Matthieu Crozet

Do migrants follow market potentials?

An estimation of a new economic geography model

Matthieu Crozet *

Abstract

New Economic Geography models describe a cumulative process of spatial agglomeration: Firms tend to cluster in locations with good access to demand, and similarly, workers are drawn to regions where market potential is high because the price index is lower there. This paper provides an empirical assessment of this forward linkage that relates labour migrations to the geography of production through real wage differentials. In the spirit of Hanson (1998), we use bilateral migration data for five European countries over the 1980s and 1990s to perform quasi-structural estimations of a new economic geography model derived from Krugman (1991). The results show strong evidence in favor of this model. As expected, migrants do follow market potential. Moreover, we provide estimates for all key parameters of the model. These estimates suggest that a sudden emergence of a core-periphery pattern is unlikely within European countries: centripetal forces are too limited in geographical scope, and mobility costs are too high.

Keywords: Agglomeration, economic geography, European regions, migration.
JEL Classification: F12, L11, R12, R23.

This version: October 2003.

1 Introduction

Over the past decade, the *new economic geography* has gained remarkable momentum. The theoretical tools derived from the seminal contributions of Krugman (1991, 1992) and Krugman and Venables (1990, 1995) have contributed to an improved analysis of economic agglomerations. They have also engendered considerable interest from researchers and policymakers, especially in the European Union, where there is concern that further integration and successive enlargements may threaten regional cohesion (Puga, 2002). Although a great deal of theoretical literature has followed this paper

*TEAM - University of Paris 1 Panthéon-Sorbonne - CNRS, 106-112 Bd de l'Hôpital - 75657 Paris Cedex 13. *email* <crozet@univ-paris1.fr>.

(See Fujita, Krugman and Venables, 1999 and Neary, 2001 for reviews of the current state of theoretical knowledge), *new economic geography* (henceforth referred to as NEG) has not yet generated a comparable volume of empirical literature. Hence, empirical work has a prominent place in the research agenda, : as propounded by Fujita *et al.* (1999)[1]: "*[...] we clearly need much more [empirical] work as closely tied to the theoretical models as possible. [...] Under what conditions do economies really spontaneously evolve a core-periphery pattern? Is Europe really going to be able to maintain its polycentric industrial geography?*" (Fujita *et al.*, 1999. p. 347-348).

The aim of this article is to help fill this gap and answer the questions raised by Fujita *et al.*. We use annual inter-regional migration data for five European countries over the 1980s and 1990s to investigate the relevance of the NEG framework. Moreover, the paper attempts to measure the strength of the centripetal forces that may affect the geography of European countries.

Krugman's (1991) original model describes a Hirschmann-type cumulative process (Hirschmann, 1958)of spatial agglomeration based on the interaction of two centripetal forces. The first force at work - backward linkage - influences the location choice of firms: in the presence of transport costs and scale economies, a region with a good access to markets is the preferred location for a firm. In the same way, forward linkage influences the location choice of individuals for central markets: in a region offering good access to a large range of commodities, the cost of living is lower because consumption incurs lower transportation costs. Put together, these two forces reinforce each other and encourage firms and workers to cluster[2].

Therefore, in a standard NEG framework, agglomeration may occur because access to markets positively influences the location choices of both firms and workers. However, most empirical investigations explicitly referring to NEG models are mainly devoted to the assessment of backward linkage. Indeed, one can roughly divide the existing empirical literature into three groups, all of which explain firm behavior through

[1]The same call for more empirical research of this nature was made by Neary (2001).

[2]This agglomeration process is balanced by a dispersion force that emanates from the greater competition that arises when the concentration of firms increases. It should be noted that input-output relations between industries may also play the role of forward linkage and induce agglomeration even when the model does not allow for migration (Krugman and Venables, 1995, Puga, 1999).

the importance of market access (see Overman, *et al.* (2003) for a comprehensive survey). First, a significant part of the empirical research aims at assessing whether greater access to demand in a region favors local firms. Such a relation, known as the "home market effect"[3], plays a large part in backward linkages. Recent studies, such as Davis and Weinstein (1999), Head and Ries (2001), or Trionfetti (2001), have found strong evidence of home market effects. Secondly, the literature on firms' location choice has supported the backward effect hypothesis. Assuming that firms choose among alternative locations in order to maximize their expected profits, these studies confirm that plants are drawn to regions with good access to demand (see, for instance, Friedman *et al.*, 1992, Devereux and Griffith, 1998, Head and Mayer, 2002, Crozet *et al.*, 2003). Thirdly, Hanson (1998) and Redding and Venables (2000) explore the spatial correlation of factor prices and demand. These two papers exploit an equilibrium equation of the NEG models, *i.e.*, the wage equation that relates nominal wages in each region to its distance from economic centers. Here again, the underlying intuition is the backward effect: a greater access to markets ensures higher profits for local firms, and thus higher wages in the long run. Using cross-country data, Redding and Venables (2000) show that good access to sources of supply and demand positively affects per capita incomes. Hanson (1998) studies the wage equation of a new economic geography model based on Krugman (1991) by performing a nonlinear estimation based on US county panel data. Such structural estimation proves the positive influence of the neighboring incomes on local wages, which is consistent with strong backward linkages among regions. Moreover, Hanson uses nonlinear least squares estimates of the model's key parameters of in simulations suggesting that agglomeration forces are limited in geographic scope. The economic influence on surrounding wages that emanates from a given county decreases rapidly with distance: it is effective only within a radius of less than 1000 kilometers.

In the spirit of Hanson (1998), we perform a quasi-structural estimation of the NEG model. However, unlike most of the empirical literature on new economic geog-

[3] "Home market effect" states that, in an increasing return industry, an increase in local demand results in a more than proportional increase in production and employment. This theoretical result is a standard prediction of monopolistic competition trade models. It also appears in NEG models.

raphy, the focus here is on forward linkage. We analyze the core equation of the NEG model that relates labour migrations across regions to the geography of production through real wage differentials. Hence, the paper examines whether access to markets (which defines the theoretical price indexes) has a significant positive influence on migration choices. Econometric results provide strong evidence for this forward linkage. Moreover, the structural NEG framework is found to have a better fit than a simpler competing model. Finally, simulations based on parameter estimates show that the geographic scope of centripetal forces is quite limited. These forces are too weak to fully balance the high barriers to migration that affect the location choices of individuals. Hence, these empirical results suggest that forward linkages and labour mobility are not sufficiently strong to cause a rapid evolution of the geography of European countries toward a core-periphery pattern.

The rest of the paper is organized as follows. The next section provides the theoretical framework for our analysis. The third section discusses some estimation issues and describes the data used. Section 4 presents econometric results for the two specifications of the migration equation: a simple gravity model and the structural model of NEG. In section 5, parameter estimates of the NEG model are used in simulation exercises. Section 6 concludes.

2 Theoretical framework

This section derives a tractable migration equation from a new economic geography model. This framework emphasizes the role of access to markets in regional dynamics. Indeed, agglomeration processes are driven by centripetal forces appearing to be closely related to Harris' (1954) market potential functions. Harris defined a measure of the demand that a firm faces as the economic size of all surrounding locations weighted for distance: $\text{Market Potential}_i = \sum_{j=1}^{R} (Y_j/d_{i,j})$, where R is the number of locations within the relevant area, Y_j the economic size of region j, and d_{ij} is the geographical distance between locations i and j $(i, j \in [1, R])$. NEG shows from a well-defined model that both firms' and workers' location choices are related to such a measure of access to markets.

Following Hanson (1998), we extend Krugman's (1991) framework, introducing a nontraded good which generates more realistic spatial dynamics. Moreover, we consider migration costs and exogenous employment opportunities.

2.1 Production and consumption

The market consists of R regions endowed with two factors: immobile and mobile labour. Each region produces three goods: A homogeneous "traditional" good (z), non-traded services (y) and manufactured goods (x).

Commodity z is assumed to be homogeneous and produced under perfect competition. It is traded costlessly across regions and employs immobile labor only. Therefore, the price of good z and the wages for immobile labour are the same everywhere. Taking the price of z as the numeraire, we have $p_z = 1$ in all regions.

Both manufactured goods and services are monopolistically competitive industries. These industries employ mobile labour to produce horizontally differentiated varieties. The production of each variety is subject to economies of scale. Within each industry, the labour required to produce a quantity q is respectively : $\beta_x q_x + \varepsilon_x$ and $\beta_y q_y + \varepsilon_y$, where β_x and ε_x (resp. β_y and ε_y) are marginal and fixed input requirements for production in industry x (resp. y). If n_{xit} and n_{yit} denote the number of varieties of good x and y produced in region i at date t, the total employment in each industry or region i at date t is:

$$L_{it}^x = n_{xit}\left(\beta_x q_{xit} + \varepsilon_x\right) \quad \text{and} \quad L_{it}^y = n_{yit}\left(\beta_y q_{yit} + \varepsilon_y\right) \quad i \in [1, R] \tag{1}$$

Consumers have identical Cobb-Douglas preferences over goods:

$$U_{it} = C_{xit}^{\mu} C_{yit}^{\phi} C_{zit}^{1-\mu-\phi}, \quad i \in [1, R] \tag{2}$$

μ, ϕ and $(1-\mu-\phi)$ are expenditure shares for manufactured goods, services and the traditional good, respectively. C_{zit} is the quantity of traditional good consumed in region i at date t. C_{xit} is a composite of manufacturing product varieties:

$$C_{xit} = \left(\sum_{m=1}^{n_{xt}} c\left(m\right)_{xit}^{\frac{\sigma_x - 1}{\sigma_x}} \right)^{\frac{\sigma_x}{\sigma_x - 1}}, \quad i \in [1, R] \tag{3}$$

Where σ_x denotes the elasticity of substitution between varieties, $c\left(m\right)_{xit}$ is the quantity consumed of variety m in region i at date t, and n_{xt} is the number of available varieties in the economy ($n_{xt} = \sum_{i=1}^{R} n_{xit}$). Consumers cannot import service varieties from other regions; therefore, the number of available y varieties in region i is the number of varieties produced within the region (n_{yit}), and C_{yit} is:

$$C_{yit} = \left(\sum_{m'=1}^{n_{yit}} c\left(m'\right)_{yit}^{\frac{\sigma_y - 1}{\sigma_y}} \right)^{\frac{\sigma_y}{\sigma_y - 1}}, \quad i \in [1, R] \tag{4}$$

As usual in this framework, all producers have the same profit-maximizing price, which is a constant markup over marginal cost. Denoting w_{it} as the mobile workers' wage in region i at date t, the *fob* price of a variety produced in region i is:

$$p_{xit} = \frac{\sigma_x}{\sigma_x - 1} \beta_x w_{it} \quad \text{and} \quad p_{yit} = \frac{\sigma_y}{\sigma_y - 1} \beta_y w_{it} \quad i \in [1, R] \tag{5}$$

Moreover, free entry in each sector leads to zero-profits at equilibrium. Therefore, using equations (1) and (5) and the equilibrium condition for each regional labour market, one can derive the number of firms in each region:

$$n_{xit} = \frac{L_{it}^X}{\varepsilon_x \sigma_x} \quad \text{and} \quad n_{yit} = \frac{L_{it}^Y}{\varepsilon_y \sigma_y} \quad i \in [1, R] \tag{6}$$

Finally, we allow for iceberg transport costs in shipping manufactured goods between regions. We assume a fraction $(\tau_{ij} - 1)/\tau_{ij}$ of the good melts away in transportation so that $\tau_{ij} > 1$ units of the good have to be exported from region i to deliver one unit to region j. This transport cost is assumed to be an increasing function of the distance between the two regions d_{ij}:

$$\tau_{ij} = B d_{ij}^{\delta} \qquad \forall i, j \in [1, R], \qquad \delta > 0 \quad \text{and} \quad B > 0 \tag{7}$$

2.2 Market potential function

Recalling that the price of the traditional good is normalized to one, the real wage of mobile workers in region i is simply:

$$\omega_{it} = \frac{w_{it}}{P_{yit}^{\phi} P_{xit}^{\mu}} \tag{8}$$

where P_{xit} (respectively P_{yit}) is the CES price index of the aggregate of industrial (resp. service) goods in region i:

$$P_{xit} = \left[\sum_{k=1}^{k=R} \left(\sum_{m=1}^{n_{xkt}} \left(B d_{ik}^{\delta} p_{xmkt} \right)^{1-\sigma_x} \right) \right]^{\frac{1}{1-\sigma_x}} = \left[\sum_{k=1}^{k=R} n_{xkt} \left(B d_{ik}^{\delta} p_{xkt} \right)^{1-\sigma_x} \right]^{\frac{1}{1-\sigma_x}} \tag{9}$$

$$P_{yit} = \left(\sum_{m'=1}^{nit} p_{ym't}^{1-\sigma_y} \right)^{\frac{1}{1-\sigma_y}} = n_{yit}^{\frac{1}{1-\sigma_y}} p_{yit} \tag{10}$$

It is clear from (9) that the price index of manufactured goods can be thought of as the inverse of a market potential function: it exhibits a comparable sum of market sizes in all regions weighted by distances. Therefore, its interpretation is straightforward. The price index is higher in remote regions where consumers have to import a large part of their demand from distant locations. Similarly, leaving aside nominal wage, workers' real income is lower in regions offering a relatively small number of service varieties. This *price index effect* makes regions with a high density of services and low-cost access to large manufacturing markets more attractive places to live. It is precisely the Hirschman-type forward linkage that contributes to the cumulative process of spatial agglomeration.

2.3 Migration choice

Our model of migration follows that of Tabuchi & Thisse (2002). We consider a mobile worker k from region j and his location choice among R regions (including j). His migration choice results from a comparison of the perceived quality of life in the various locations[4]. For empirical convenience, we assume that the migration decision is designed to maximize the following objective function:

$$\pi^k_{ji,t} = V^k_{ji,t} + \epsilon^k_i = \ln\left[\omega_{i,t}\rho_{i,t-1}\left[d_{ij}(1+bF_{ij})\right]^{-\lambda}\right] + \epsilon^k_i \qquad i \in [1, R] \tag{11}$$

where $\rho_{i,t}$ is the employment probability for an immigrant in region i at date t and $\left[d_{ij}(1+bF_{ij})\right]^{\lambda}$ is a migration cost which increases with the distance between home and host regions. λ and b are strictly positive coefficients, and F_{ij} is a dummy variable indicating whether regions i and j *do not* share a common border. ϵ^k_i is a stochastic component capturing k's personal perception of the characteristics of region i. To avoid an endogeneity problem when turning to empirical application, we assume that migration choices at date t are determined from a comparison of V^k_{ji} across regions at date $t-1$. Therefore, individual k will choose to locate in region i if $V^k_{ji,t-1} > V^k_{jr,t-1}, \quad \forall \quad r \neq i$. With these convenient assumptions on distribution of ϵ^k_i, the probability of choosing region i is given by the logit function:

$$P(M_{ji,t}) = e^{V^k_{ji,t-1}} \Big/ \sum_{r=1}^{R} e^{V^k_{jr,t-1}} \tag{12}$$

The expected migration flow between regions j and i is $L_{j,t}P(M_{ji,t})$. Noting that, similarly, the total outflow from j is $L_{j,t}\left[1 - P(M_{jj,t})\right]$, the share of emigrants from region j choosing to go to region i is:

$$\frac{migr_{jit}}{\sum\limits_{i' \neq j} migr_{ji't}} = \frac{e^{V^k_{ji,t-1}}}{\sum_{r=1}^{R} e^{V^k_{jr,t-1}} - e^{V^k_{jj,t-1}}}$$

Using equations (5), (6), (8) (9), (10) and the definition of $V^k_{ji,t}$, this share can be

[4]Following standard NEG models, we make the strong simplifying assumption that migration decisions are based only on migration costs and current expected real wage differences. However, Baldwin (2001) shows that allowing for forward-looking expectation does not significantly change the theoretical conclusions when migrations costs are sufficiently high.

written as:

$$\ln\left(\frac{migr_{jit}}{\sum\limits_{i'\neq j} migr_{ji't}}\right) = \ln\left[\left(L^{Y}_{i(t-1)}\right)^{\frac{\phi}{\sigma_y-1}}\right] + \ln\left[\left(\sum_{k=1}^{k=R} L^{X}_{k(t-1)}\left(w_{k(t-1)}d^{\delta}_{ij}\right)^{1-\sigma_x}\right)^{\frac{\mu}{\sigma_x-1}}\right]$$
$$+ \ln\left[w^{1-\phi}_{i(t-1)}\rho_{i(t-1)}\right] + \ln\left[d_{ij}\left(1+bF_{ij}\right)\right]^{-\lambda} + \tilde{a}_{j(t-1)} \qquad (13)$$

With $\tilde{a}_{j(t-1)} = -\ln\left(\sum_{r=1}^{R} e^{V^{k}_{jr,t-1}} - e^{V^{k}_{jj,t-1}}\right)$.

Equation (13) captures the trade-off faced by potential migrants who have to choose among several possible locations. The left-hand side of equation (13) is the share of migrants from a given region who have decided to move to region i. On the right-hand side, the third term represents the expected wage in the region, which increases with the host region's nominal wage and the probability of being employed in this region. The fourth term captures the impact of bilateral distance on migration flows and is interpreted as a measure of mobility cost. The first two first terms denote region i's access to markets: they are, respectively, the price indices for non-traded service varieties and for manufactured goods in region i. The second term of equation (13) is clearly the most important term in this equation. It corresponds to a market potential function and relates labour migrations to the location of industrial activities and can therefore be seen as the forward linkage emphasized by NEG. Moreover, the main parameters of the NEG framework (elasticity of substitution and parameters of the trade cost function) can be estimated from this price index function. Hence, if the empirical analysis confirms that this price index actually governs migration flows, i.e. that *migrants do follow market potentials*, it will validate the role of forward linkage as a part of the endogenous agglomeration process.

3 Econometric specifications and data

We estimate two specifications of equation (13).

Clearly, equation (13) is closely related to a simple gravity equation. Besides nominal wages and employment probability, the migration flow between two regions increases with the size of the host region and decreases with the geographic distance between the two locations. Such a relation, while it is only a reduced form of equation (13), may provide a good starting point for assessing whether migrants, as firms, are attracted to large regional markets[5]. Such a gravity equation is a benchmark which allows for the identification of possible specification issues and provides a grounded competing model to the complete NEG framework.

There are several issues to address before performing estimations. An obvious proxy for the probability of finding a job in the host region would be the regional employment rate $E_{i(t-1)}$ (*i.e.* one minus unemployment rate), which is of course correlated with nominal wages. Hence, to circumvent multicollinearity problems, we consider the *expected nominal wage* as a single variable defined by the product of nominal wage and employment rate: $probw_{i(t-1)} = w_{i(t-1)}E_{i(t-1)}$. Furthermore, variables $\tilde{a}_{j(t-1)}$ do not depend on destination region i. We thus allow for a more robust specification replacing $\tilde{a}_{jt}$ with a time trend and fixed effects relative to home regions. Referring to equation (13), we expect these fixed effects to be negative. Moreover, we introduce the logarithm of the area of host region ($\log(S_i)$) in order to control for the bias resulting from the inclusion of unequally-sized regions in the sample[6]. In order to control for peculiar structural difficulties of possible host regions, we also introduce a dummy variable set to 1 for host regions that are eligible for the European Commission regional funds given under Objectives 1 or 2 (obj_i)[7]. Finally, the gravity equation to be estimated is:

$$\log\left(\frac{migr_{jit}}{\sum\limits_{i'\neq j} migr_{ji't}}\right) = \beta_1 \log\left(L_{i(t-1)}\right) + \beta_2 \log(probw_{i(t-1)}) + \beta_3 \log(d_{ij}) + \beta_4 F_{ij} + \beta_5 \log(S_i) + a_j + \beta_6 \log(trend) + obj_i + v_{ijt}, \tag{14}$$

where $L_{i(t-1)}$ is total employment in region i, a_j is a full set of home region fixed effects standing in for variables $\tilde{a}_{j(t-1)}$ in equation (13), and v_{ijt} is an error term.

[5] See for instance Helliwell (1997) for a similar study focusing on migrations in North America.

[6] This bias may be important: For instance, the largest German region in our sample (Bavaria - 70554 km^2) is more than 170 times bigger than the smallest one (Bremen - 404 km^2).

[7] Objective 1 promotes the development of regions with GDP per head below 75% of the Union average, and Objective 2 is aimed at promoting the conversion of areas affected by industrial decline.

Regional attractiveness should increase with the expected wage and decrease with distance from the source region. Above all, one should expect coefficient β_1 to be significantly positive, since the NEG framework suggests that larger regions offer a lower cost of living. Moreover, the model presented in the previous section suggests that sectoral components of gross regional product have different influences on migration choices. Thus, we also estimate a gravity equation in which regional employment is split up into three industries (services Y, manufactured goods X and agriculture Z).

The second specification to be estimated is directly taken from the theoretical model. Introducing $probw_{i(t-1)}$ and a_j into equation (13), we obtain the following non-linear testable equation:

$$\begin{aligned}\log\left(\frac{Migr_{ji,t}}{\sum\limits_{k\neq j} Migr_{jk,t}}\right) &= \frac{\mu}{\sigma_x - 1}\log\left(\sum_{k=1}^{k=R} L^x_{k(t-1)}\left(w_{k(t-1)}\left(d_{ik}\right)^{\delta}\right)^{1-\sigma_x}\right) \\ &+\alpha_1 \log\left(L^y_{i(t-1)}\right) + \alpha_2 \log\left(probw_{i(t-1)}\right) - \lambda \log\left(d_{ij}\left(1 + bF_{ij}\right)\right) \\ &+\alpha_3 \log\left(S_i\right) + a_j + \alpha_4 \log\left(trend\right) + obj_i + u_{ijt}\end{aligned} \quad (15)$$

The differences between the gravity equation (14) and equation (15) are twofold. First, the latter specification controls not only for the attraction of the local supply of manufactured goods but also for access to all surrounding markets. Secondly, this specification provides estimates of key parameters of the NEG framework (σ_x and δ). Note that these parameters are related to the supply side of the model; thus, obtaining consistent estimates for these parameters from migration flow data would indicate the relevance of the cumulative process of agglomeration featured by NEG.

A major difficulty with equation (15) arises from the definition of the traditional sector. According to the theoretical framework, the difference between sector x ('manufactured goods') and z ('traditional good') lies in market structure and the presence of scale economies: the 'traditional' sector should stand for all homogeneous productions with constant returns to scale, while all tradable and differentiated productions with increasing returns to scale should be considered as 'manufactured goods'. Unfortunately, we do not have detailed sectoral data at the regional level allowing such a classification. The simplest solution, therefore, is to consider agriculture as a proxy for 'traditional' production, so that the x sector stands for all manufactured goods (model 1). To test the robustness of the results, we also perform regressions considering both manufactured and agricultural goods as belonging to the x sector. In this specification (refereed to as model 2), $L^x_{i(t-1)}$ becomes the sum of industrial and agriculture employment, $w^x_{i(t-1)}$ is the mean wage in agriculture and industry, and $probw_{i(t-1)}$ is the product of employment rate and the mean regional wage. For both models 1 and 2, the parameters to estimate are the same. They are displayed in table 1.

Insert Table 1 here

We perform estimations of equations (14) and (15) for five European countries (Germany, Spain, The Netherlands, Italy and the United Kingdom). The data required, excluding distances, are available from *Regio* database (Eurostat).

Migration data

Regio provides annual bilateral migration data at the regional level. While precious, this information has three main shortcomings. First, data is limited to internal migrations, and we have no information (at the regional level) about migrations among European countries. Second, *Regio* does not provide data at a very detailed geographic level: data is available at the NUTS 1 or NUTS 2[8], depending on the country (see data appendix). Finally, the time span is not exactly the same for all of the countries: 1983-1992 for Germany, 1983-1993 for Spain and Italy, 1988-1994 for the Netherlands and 1980-1985 Great Britain. For these reasons, we do not pool all the data together but perform separate regressions for each country[9].

[8]NUTS (*nomenclature of territorial units for statistics*) is a nomenclature providing a hierarchical structure of sub-national regions covering all European territory. Eurostat first subdivides countries of the EU into 78 NUTS 1 (corresponding for instance to German Landers), each of them being divided into several NUTS 2.

[9]Indeed, the higher costs of cross-border migrations lead to a decreased relevance of international flows relative to internal ones (Helliwell, 1997).

Market size and expected wages

Regio also provides data on sectoral employment, wages, unemployment rates and areas at a regional level. Available British regional unemployment data only starts in 1983 so that keeping employment rates in the estimations would drastically reduce the number of observations. Hence, in the case of Britain, the variable $probw_{i(t-1)}$ is simply taken to be the product of the nominal wage for each year $(t-1)$ and the employment rate in 1983.

Distances

The model also requires a grounded measure of bilateral distances between all regions of the same country ($d_{ij}\ \forall i \neq j$). This variable, which stands for both trade costs and migration costs, is of particular importance. We greatly simplified measurement issues, dropping data involving overseas territories, islands and Ulster, so that the distance between two regions may be proxied by road distances between their respective capital cities. Distances are estimations provided by an electronic road atlas [10] that calculates the length of the quickest route between the two cities[11]. Thus, this measure takes into consideration geographic elements such as mountains, lakes, density and quality of road infrastructures... Besides inter-regional distances ($d_{ij},\ i \neq j$), one needs a proxy for internal distances d_{ii}. Indeed, the market potential term of equation (15) includes, as for goods imported from abroad, the transport cost charged on x goods produced and consumed locally. The internal distance is proxied by: $d_{ii} = (2/3)\sqrt{S_i/\pi}$, where S_i denotes the area of the region (Redding and Venables, 2002).

[10] *ROUTE 66 Geographic Information Systems.*

[11] These distances are available on line: `http://team.univ-paris1.fr/trombi/crozet/data.htm`

4 Results

4.1 The gravity equation

Equation (14) is estimated by ordinary least squares with a full set of fixed effects for source regions. Tables 2 and 3 display regression results[12].

Insert Table 2 about here

Insert Table 3 about here

The gravity-type model has a fairly high explanatory power. R^2 are high, most coefficients have the expected sign, and estimates from different countries have comparable magnitudes. Distance always has the expected negative influence on migration. The high value of this coefficient shows how reluctant European workers are to move to a distant region. Moreover, migration costs seem to be even higher for long-distance migrations: crossing more than one regional border reduces migration flows, at the leastby 33% for Great-Britain ($i.e\ \ 1 - 1/\exp(0.402)$) and at the most by 55.5% for Italy ($i.e\ \ 1 - 1/\exp(0.809)$).

However, $probw_{i(t-1)}$ does not perform well. For Great Britain, the coefficient on this variable is positive, as expected, but it is negative for the other countries. In the case of Germany and the Netherlands, this result is probably the consequence of the very low variability of wage and employment rates across regions. Indeed, a positive collinearity for $probw_{i(t-1)}$, fixed effects a_j and the trend variable shed doubt on the robustness of this estimation. Multicollinearity among these regressors probably also explains why coefficients on a_j are positive, whereas the theoretical framework suggests they should be negative. We confirm this intuition, regressing for these countries a model excluding the variable $probw_{i(t-1)}$. It shows very positive properties: multicollinearity disappears and we obtain negative values for variables a_j. Results on other variables do not change significantly. For the other countries, counter-intuitive

[12] To save space, we report neither the coefficients on the dummies nor those on the fixed effects. All results are, of course, available upon request.

results on variable $probw_{i(t-1)}$ mainly reflect an issue commonly encountered in empirical studies of internal migrations if one does not control for the migrant's individual characteristics and regional labour market structures[13]. For instance, in countries where regional inequalities are substantial, such as Italy, one expects very important centripetal migrations. However, differences in regional industrial structure reduce the opportunities for workers from remote (low wage) regions to find jobs in a central (high wage) location (Faini *et al.*, 1997). Hence, in the case of Italy, we performed regressions without considering migrations toward southern regions[14]. The focus on migrations toward leading regions provides more convincing results (Cf. columns *Italy-Total* of table 3).

Finally, we observe a positive influence of total employment in the host region. This significant effect confirms that migration patterns reveal centripetal dynamics. However, considering separately the influence of employment in services, manufacturing and agriculture, it appears that the positive influence of economic size is mainly due to services. Local manufacturing employment has very little influence on regional attractiveness, and its influence is even significantly negative for Germany and Spain. Interpretation of this result is twofold. On the one hand, it may suggest that spatial distribution of manufacturers does not influence migrants' location choices. Hence, workers would not move for better access to manufactures, as is suggested by the price index effect at the heart of NEG models. On the other hand, it may suggest that local employment in the tradable good is not an relevant proxy for regional access to markets. Indeed, manufacturing firms supply all locations, so that local production is only a slight part of access to markets defining the local price index and the incentives to migrate. Thus, this result justifies the use of a real market potential function in the spirit of NEG framework. These two interpretations lead to opposite conclusions about the relevance of the NEG framework. The structural estimation of equation (13) should settle this issue.

[13] See for instance, Greenwood 1975, Herzog *et al.* 1993 and Borjas 1999 for reviews and Decressin and Fatàs, 1995, and Faini *et al.*, 1997 for studies of the weak responsiveness of geographic mobility to expected income differentials within European countries.

[14] We drop Lazio, Abruzzo-Molise, Campania and Puglia as possible destination regions.

4.2 The NEG framework

We now turn to the estimation of equation (15) by non linear least squares. Obviously, it is not possible to provide an estimation of both μ and σ_x since μ is not independently attached to a variable. A simple way to overcome this problem is to treat μ as an exogenous parameter. Recalling that μ is the expenditure share of x goods, we impose $\mu = 0.4$ in model 1 (where x stands for manufactured goods only) and $\mu = 0.6$ in model 2 (where x represents both agriculture and manufacturing)[15]. Tables (4) and (5) display our estimation results. We do not report the coefficients on fixed effects a_j, but, in accordance with theoretical predictions, they proved to be significantly negative in all regressions[16].

Insert Table (4) about here

Insert Table (5) about here

Most parameters, in particular those which define the price index, converge toward consistent values. The two specifications of (15) lead to comparable estimations. The results show substantial differences among the various countries, but the coefficients always have the same order of magnitude. Moreover, the results for equation (15) are very similar to those estimated in the previous section. Following Hanson (1998), we calculate the Schwarz information criterion to determine if this model is more appropriate than the simple gravity equation[17]. In most cases (with the exception of Italy), the Schwarz criterion indicates that the structural NEG model provides a better fit.

[15] As a robustness test, we experimented with alternative values for μ. Higher values for μ lead to higher estimates of σ_x and δ, but the other parameters remain almost unchanged.

[16] As in the previous section, fixed effects are positive for The Netherlands when $probw_{i(t-1)}$ remains in the model.

[17] The Schwarz - or Bayesian - information criterion provides a simple method for model selection. Defining M to be the maximum value of the likelihood, n the number of observations and k the number of parameters, the criterion proposes to select the model for which $-2\log M + k\log n$ is smallest.

Migration costs and expected wage

The order of magnitude of the estimated coefficients on expected wage (α_2) is very close to the one obtained by the gravity-type equation. Results also show a great influence of the variables capturing migration costs (λ and b) over migration flows. Compared to the other Europeans, the British seem to be the most willing to move. A small coefficient on distance (λ) is also observed for Italy. However, results for this country also exhibit a troublingly high value for the b coefficient. Here again, estimations on the restricted sample lead to more consistent results (see colons *Italy-North* in table 5).

Elasticities of substitution and transport cost: Migrants do follow market potentials

In contrast to the simple gravity function, this second specification provides an estimation of the parameters of the structural market potential function. We see from tables (4) and (5) that this structural model generates very encouraging results: All parameters defining the CES price indices (σ_x, δ and α_1) have the expected sign and are highly significant.

Results show, as expected, a strong positive influence of service employment in the host region on migration flows. Coefficients α_1 converge toward values between 0.43 and 0.97. To interpret this result, it is useful to recall that the theoretical model allows for the derivation from coefficients α_1 of the estimated values of the elasticity of substitution between any pair of service varieties: $\sigma_y = \frac{\phi}{\alpha_1} - 1$. According to the values chosen for μ, we set ϕ to 0.4, which is a reasonable value for the expenditure share of non-traded services. Thus, computed values of σ_y lie between 1.41 for Italy to 1.93 for the Netherlands. These estimations are consistent with the constraints imposed by the theory. They are relatively low, which denotes a high degree of product differentiation.

Whereas gravity-type analysis showed that the size of local manufacturing employment in the destination region does not significantly affect migration decisions, we observe here that parameters defining the market potential function are all significant.

In accordance with the NEG model's prediction, access to manufactured commodities do influence workers' mobility since it is measured by a grounded market potential function.

Indeed, elasticities of substitution between manufacturing varieties (σ_x) are always strictly positive, and, as required by theory, significantly superior to one in every country. They vary between 1.3 (UK) and 4.3 (The Netherlands). Aggregating both manufacturing and agriculture in the x sector, elasticities take higher values (from 1.5 to 5.6) but remain very close to those derived from the first specification. These values of elasticities of substitution are lower than those observed by Hanson (1998) and in other related recent studies such as Baier & Bergstrand (2001) and Head & Ries (2001)[18].

One may be surprised by the substantial difference between our estimates of σ_x and Hanson's results. However, in our specification, the estimated value of σ_x is directly dependent on the chosen value of μ. A higher value for μ should have led to higher σ_x. Hanson's estimations of μ are close to 0.9; hence, the ratios $\frac{\mu}{\sigma-1}$ he estimates range between 0.13 and 0.25, which is comparable to the estimates reported in tables (4) and (5).

Estimated values of δ are always strictly positive but vary considerably between countries: They are greater than 3 for Germany and the Netherlands, but only around 0.5 for Spain. Consequently, the variation across countries of the complete coefficient on transport cost ($\delta(1-\sigma_x)$) is important (more than 10 for Germany and less than 0.5 for Spain and Great Britain). A high $\delta(1-\sigma_x)$ is a sign that the economic activity in surrounding regions have little influence on regional price indexes. In other words, this suggests that regional market potentials (as perceived by migrants) are essentially influenced by the local level of production. The low $\delta(1-\sigma_x)$ observed for Spain and United Kingdom can be explained by a greater sensibility of migrants to the differences in market access, or the higher level of regional specialization in these countries, which involves a greater influence of the market potential function in migration decisions.

[18] Hanson's estimates of market potential lead to elasticities of substitution between 5 and 7.6. Head & Ries (2001) and Baier & Bergstrand (2001) both study international trade flows ; their estimates of σ_x lie between 6 and 11.

5 Simulation exercices

This structural estimation provides empirical support for Krugman's model (1991). It also makes it possible to go further and evaluate the scope of agglomeration dynamics at stake within each of the five European countries. Indeed, the strengths of centripetal and centrifugal forces described by the theory are defined by the parameters estimated in the previous section. To interpret the estimates in terms of spatial dynamics, this section presents two complementary numerical evaluations.

The break point

In a simple Krugman (1991) model, the analysis of the equilibrium often goes through a calculation of the conditions under which a core-periphery pattern is the only stable solution (Fujita et al., 1999). These conditions define a threshold level of inter-regional transport cost at which the economy converges towards a core-periphery pattern. This transport cost, known as the *break point* of the model, can be interpreted unambiguously as a measure of the relative scope of centripetal forces. The expression for the break point value only depends on the key parameters that define the price index (i.e. the share of expenditure on differentiated goods and the elasticity of substitution):

$$\tau^{Break} = \left[\frac{[\sigma_x(1+\mu)-1]\,(1+\mu)}{[\sigma_x(1-\mu)-1]\,(1-\mu)} \right]^{1/(\sigma_x-1)}$$

This break point is derived from a two-region model that assumes zerono intra-regional trade costs. The introduction of internal trade costs does not significantly change the model, but the break point turns out to be a threshold *relative* trade cost: $(\tau_{ij}/\tau_{ii})^{Break}$. Using (7), we can express the break point as a relative distance:

$$\left(\frac{d_{ij}}{d_{ii}}\right)^{Break} = \left[\frac{[\sigma_x(1+\mu)-1]\,(1+\mu)}{[\sigma_x(1-\mu)-1]\,(1-\mu)} \right]^{1/[\delta(\sigma_x-1)]} \qquad (16)$$

Insert Table (6) about here

Table (6) reports these minimal distances calculated with the parameters estimated in the previous section[19].

These relative distances indicate how far the agglomeration forces emanating from a region extend across space: multiplying the relative distance by the internal distance of a central region, one obtains the radius where any activity with increasing return to scale should be attracted to the core. In Germany, Italy and the Netherlands the threshold relative distances are relatively short (between 1.2 and 2.5), which suggests that the scope of centripetal forces is small. Therefore, in those countries a significant center-periphery pattern can only emerge between very close regions. For instance, the Italian region of Lombardia has an internal distance of 58 kilometers, so centripetal forces emanating from these regions should dominate within a radius around its center ranging between 94.5 km (1.63 x 58) and 150.2 km (2.49 x 58). These distances are too small to threaten neighboring regions[20]. For Germany, the magnitude of centripetal forces seems even smaller: the Bavarian region has an internal distance of about 100 km and possibly attracts activities located within a radius of only 120 to 135 km from its center. In the cases of Spain and Great Britain, the break points cannot be calculated: the *no black hole condition* that permits the existence of a dispersed equilibrium does not hold[21]. Therefore, in these two countries, workers' movements reveal extremely strong self-sustaining processes of agglomeration.

Predicted migration flows

The break point is a simple and practical tool. It is, however, limited since it does not include migration costs. Therefore, we have to compute more accurate spatial relations to appraise the real scope of agglomeration dynamics within European countries. To this end, we figure *the predicted gross share of emigrants for different relative regional sizes*. Such a simulation shows, using all estimated parameters, the strength of the response of the European workforce facing a given regional inequality.

To simplify the simulation exercise, we consider a two-region economy (i and j) with

[19]For Germany, Italy and the Netherlands, the parameters from the restricted regressions are used.
[20]Milano, the main Lombardian city, is within 141 km of Turin, 164 km of Genova, and 576 km of Roma.
[21]The *no black hole condition* is $\frac{\sigma_x - 1}{\sigma_x} > \mu$ (Fujita et al., 1999).

identical internal distances, wages, employment opportunities and shares of services in total employment. The two regions differ only in population size: Region j is assumed to be a peripheral region, so that $L_j < L_i$. Very simple manipulations of equations (12) and (11) yield the following equation for the share of emigrants in the total population of region j:

$$\frac{Migr_{ij}}{L_j} = \frac{\Delta}{\Delta + [d_{ij}(1+b)]^{\lambda}} \quad ; \quad \Delta = \frac{L_i^{\alpha_1}\left(L_i(d_{ii}^{\delta(1-\sigma)}) + L_j(d_{ij}^{\delta(1-\sigma)})\right)^{\mu/(\sigma-1)}}{L_j^{\alpha_1}\left(L_j(d_{jj}^{\delta(1-\sigma)}) + L_i(d_{ij}^{\delta(1-\sigma)})\right)^{\mu/(\sigma-1)}}$$

Figure (1) plots these emigration shares, using the parameters of model 1 displayed in tables (4) and (5)[22]. Internal distances d_{jj} and d_{ii} are set to 75 km, which is roughly consistent with the real size of large European regions.

Insert Figure1 here

As expected, the number of workers who choose to move to the core region is greater when regional inequality is sizeable. The simulations underscore the weakness of the response of labour markets to regional inequalities. Excepting Great Britain, the gross emigration shares are very low, even when regional inequality is rather large[23]. Even more interesting is the finding that centripetal migration decreases with inter-regional distance in all of the five countries. Indeed, in theory the influence of inter-regional distance on migration flows results from a trade-off. On the one hand, the model suggests that a greater inter-regional distance strengthens the forward linkage that contributes to the agglomeration dynamics: a higher transportation cost increases the cost of living far away from the larger market and thus heightens migration incentives. On the other hand, inter-regional distance positively influences migration costs. Hence,

[22] As previously, we use the estimates from the restricted regressions for Germany, Italy and the Netherlands.

[23] Note that these values are all slightly larger than but roughly consistent with the real mean gross migration shares.

the overall negative influence of distance on migrations provides evidence that workers' sensitivity to migration costs tends to overcome agglomeration forces. In other words, labour mobility in Europe is sufficiently low to make the swift emergence of a core-periphery pattern very unlikely at a large geographical level. This result applies for all of the five countries, even those for which the magnitude of centripetal forces is rather large, as suggested by the break point analysis. For instance, in the Netherlands, table (6) suggests that a core-periphery pattern may be sustainable for a relative inter-regional distance of 2. However, figure (1) shows that, for such a distance, less than 1% of the population in the periphery will move to the core region. The British case is also interesting. Just as for the other countries, the relation between distance and migrations is negative. However, the slope is smaller and the estimated shares of emigrants are distinctly larger. Hence, the geographic scope of forward linkage is much larger in this country.

6 Conclusion

This paper aims to contribute to the empirical assessment of new economic geography models. Recent developments in spatial agglomeration theories, based on Krugman's (1991) article, are abundant and have received particular attention from policymakers, especially in Europe. The basic intuition of the model highlights the influence of access to markets - represented by a market potential function - on location choices of both firms and workers. The cumulative process of agglomeration rests on the complementarity of these two relations: agglomeration may occur only if migrants, like firms, are attracted by high market potential regions.

Although most empirical papers on the field focus on the backward effect (which relates the location choices of firms to the geography of demand), we use interregional migration data for five European countries to establish the validity of forward linkage (which relates the location choices of individuals to the geography of supply).

First, a gravity equation provides a benchmark analysis of the determinants of labour migration within European countries. The results indicate that wealthy regions attract more migrants. However, this is mainly due to the influence of the local supply of services, whereas local manufacturing employment seems to have no influence on migration flows. This result may be interpreted either as a failure of the *price index effect* hypothesis, or as a justification for the use of a more grounded model. Thus, the paper presents structural estimates of the migration equation derived from a NEG model. Such estimation yields to the core results of this study.

Forward linkage emphasized by New Economic Geography models is relevant. Access to markets positively influences migration choices through the industrial price index effect. The relatively good fit displayed by the New Economic Geography model (compared to a simple gravity model) and the concordance in sign and magnitude between the estimated parameters and the theoretical predictions proves the empirical validity of this theoretical framework. Finally, the parameter estimates allow for evaluation of the extent of agglomeration dynamics at work within European countries. Simulations of the theoretical model suggest that centripetal forces - except in Spain and Great Britain - are very limited in geographic scope. Moreover, we show that, in all of the five countries, barriers to migration are high enough to balance the centripetal forces. Thereby, partly because of the low propensity to migrate, it seems very unlikely that a catastrophic core-periphery pattern will emerge within European countries, or *a fortiori* on a greater scale.

Acknowledgments

I am grateful to Pierre-Philippe Combes, Maria Crawford, Lionel Fontagné, Thierry Mayer, Daniel Mirza, Federico Trionfetti and three anonymous referees for a number of suggestions. I have also greatly benefited from discussions with participants in seminars at the Universities of Paris I, Paris IX, Toulouse I and Montpellier I.

References

Baier, S. L., Bergstrand, J. H. (2001) The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity. *Journal of International Economics*, 53: 1-27.

Baldwin, R. E. (2001) Core-periphery model with forward-looking expectations. *Regional Science and Urban Economics*, 31(1): 21-49.

Borjas, G. J. (1999) The Economic Analysis of Immigration. In O. Ashenfelter and D. Card (eds), *Handbook of Labour Economics*, Volume 3A. Amsterdam: North-Holland, 1697-1760.

Crozet, M., Mayer, T., Mucchielli, J-L. (2003) How Do Firms Agglomerate? A Study of FDI in France. Discussion Paper 3873, Center for Economic Policy Research.

Davis, D. R., Weinstein, D. E. (1999) Economic Geography and Regional Production Structure: An Empirical Investigation. *European Economic Review*, 43: 379-407.

Decressin, J., Fatàs, A. (1995) Regional Labour Market Dynamics in Europe. *European Economic Review*, 39(9): 1627-1655.

Devereux, M., Griffith, R., (1998) Taxes and the Location of Production : Evidence from a Panel of US Multinationals. *Journal of Public Economics* 68(3): 335-367.

Faini, R., Galli, G., Gennari, P. Rossi, F. (1997) An Empirical Puzzle: Falling Migration and Growing Unemployment Differentials among Italian Regions. *European Economic Review*, 41: 571-579.

Friedman, J. , Gerlowski, D., Silberman, J. (1992) What Attracts Foreign Multinational Corporation? *Journal of Regional Science*, 32(4): 403-418.

Fujita, M., Krugman, P. R., Venables, A. J. (1999) *The Spatial Economy: Cities, Regions and International Trade*. Cambridge: MIT Press.

Greenwood, M. J. (1975) Research on International Migration in the United States: A Survey. *Journal of Economic Literature*, 13: 397-433.

Hanson, G. H. (1998) Market Potential, Increasing Returns, and Geographic Concentration. Working Paper 6429, National Bureau of Economic Research. Revised version: mimeo December 2001.

Harris, C. D. (1954) The Market as a Factor in The Localization of Industry in the United States. *Annals of the Association of American Geographer*, 44: 315-348.

Head, K., Mayer, T. (2002) Market Potential and the Location of Japanese Investment in the European Union. Discussion Paper 3455, Center for Economic Policy Research.

Head, K., Ries, J. (2001) Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of U.S. Canada Trade. *American Economic Review*, 91(4): 858-876.

Helliwell, J. F. (1997) National Borders, Trade and Migration. *Pacific Economic Review*, 2(3): 165-185.

Herzog, H. W., Schlottmann, A. M., Boehm, T. P. (1993) Migration as Spatial Job-Search: A Survey of Empirical Findings. *Regional Studies*, 27(4): 327-340.

Hirschman, A. R. (1958) *The Strategy of Economic Development*. New Haven: Yale University Press.

Krugman, P. R. (1991) Increasing returns and economic geography. *Journal of Political Economy*, 99: 483-499.

Krugman, P. R. (1992) A Dynamic Spatial Model. Working Paper 4219, National Bureau of Economic Research.

Krugman, P. R., Venables, A. J. (1995) Globalization and the Inequality of Nations. *The Quarterly Journal of Economics*, 110(4): 859-880.

Neary, P. (2001) Of hype and hyperbolas: Introducing the new economic geography. *Journal of Economic Literature*, 39(2): 536-561.

Puga, D. (1999) The rise and fall of regional inequalities. *European Economic Review*, 43(2): 303-334.

Puga, D. (2002) European regional policy in light of recent location theories. *Journal of Economic Geography*, 2(4): 372-406.

Redding, S., Venables, A.J. (2000) Economic Geography and International Inequality. Discussion Paper 2568, Center for Economic Policy Research.

Tabuchi, T., Thisse, J.-F. (2002) Taste heterogeneity, labour mobility and economic geography. *Journal of Development Economics*, 69: 155 177.

Trionfetti, F. (2001) Using Home-Biased Demand to Test Trade Theories. *Weltwirtschaftliches Archiv*, 137(3): 404426.

A Data appendix

- **Germany. Period: 1983-1992. 10 regions (NUTS 1):** Baden-Wurttenberg, Bayern, Bremen, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Schleswig-Holstein.
- **Italy. Period: 1983-1993. 18 regions (NUTS 2):** Piemonte, Valle d'Aoste, Liguria, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria.
- **Spain. Period: 1983-1993. 15 regions (NUTS 2):** Galicia, Asturias, Cantabria, Pais Vasco, Navarra, Rioja, Aragon, Madrid, Castilla-Leon, Castilla-La Mancha, Extremaduras, Catalunia, Comunidad Valenciana, Andalucia, Murcia.
- **Netherland. Period: 1988-1994. 10 regions (NUTS 2):** Groningen, Friesland, Drenthe, Overijssel, Gelderland, Flevoland, Utrech, Noord-Holland, Zuid-Holland, Zeeland.
- **Great Britain. Period: 1980-1985. 10 regions (NUTS 1):** North, Yorkshire and Humberside, East Midlands, East Anglia, South East, South West, West Midlands, North West, Wales, Scotland.

B Tables and figues

Table 1: Coefficients and expected values

Coefficients		Expected values
σ_x	Elasticity of substitution - sector x	$1 < \sigma_x$
μ	Share of consumer expenditure on good x	$\mu = 0.4$ or 0.6
δ	Elasticity of trade costs to distance	$0 < \delta$
$\alpha_1 = \frac{\phi}{\sigma_y - 1}$	Influence of local services supply	$0 < \alpha_1 < 1$
α_3	Influence of expected wage	$0 < \alpha_3$
λ	Distance elasticity of migration cost	$0 < \lambda$
b	Influence of borders on migration cost	$0 < b$
α_4	Influence of the size of host region	$0 < \alpha_4$
a_j	Home regions fixed effects	$0 > a_j$

Table 2: Gravity-type equation (eq. 14) - OLS / fixed effects

	Dependant variable: $\log(migr_{jit}/\sum_{i'\neq j} migr_{ji't})$							
	Germany		Germany		Spain		Great Britain	
Total employment	0.934^a	–	0.735^a	–	0.944^a	–	0.794^a	–
log $L_{i(t-1)}$	(0.039)		(0.032)		(0.013)		(0.020)	
Service employment	–	0.818^a	–	0.788^a	–	1.039^a	–	0.718^a
log $L^y_{i(t-1)}$		(0.104)		(0.104)		(0.041)		(0.076)
Indus. employment	–	-0.137	–	-0.210^b	–	-0.168^a	–	0.055
log $L^x_{i(t-1)}$		(0.093)		(0.090)		(0.042)		(0.086)
Agri. employment	–	0.264^a	–	0.302	–	0.043^a	–	-0.031
log $L^z_{i(t-1)}$		(0.045)		(0.043)		(0.020)		(0.064)
Prob-wage	$-2.708^{a\S}$	$-1.126^{a\S}$	–	–	-0.357^a	-0.239^a	0.796^a	0.754^a
log $w_{i(t-1)}E_{i(t-1)}$	(0.325)	(0.387)			(0.039)	(0.042)	(0.052)	(0.109)
Distance	-0.919^a	-0.907^a	-0.908^a	-0.906^a	-0.807^a	-0.793^a	-0.454^a	-0.489^a
$\log(d_{ij})$	(0.041)	(0.041)	(0.043)	(0.041)	(0.035)	(0.035)	(0.035)	(0.034)
No-border	-0.667^a	-0.602^a	-0.643^a	-0.586^a	$-0,648^a$	-0.671^a	-0.417^a	-0.402^a
F_{ij}	(0.052)	(0.051)	(0.054)	(0.051)	(0.044)	(0.047)	(0.035)	(0.035)
Surface	-0.017	-0.014	0.115^a	0.003	0.043^a	-0.005	0.025	0.020
$\log(S_i)$	(0.024)	(0.037)	(0.018)	(0.033)	(0.013)	(0.017)	(0.022)	(0.033)
a_j (sign)	(+)	(+)	(-)	(-)	(-)	(-)	(-)	(-)
Nb. Obs.	900	900	900	900	2310	2310	540	540
R2	0.9788	0.9799	0.9772	0.9792	0.9789	0.9798	0.9905	0.9907
F-Stat	2402	2143	2365	2241	4619	4437	3205	2914
Schwarz Criterion	1238.7	1101.6	1300.2	1103.4	3564.1	3404.2	166.7	168.7

Standards errors in parenthesis; a, b = significance at 1 and 5% levels.

§: significant multicollinearity affects the fixed effect and the coefficients relating to Prob-wage.

Table 3: Gravity-type equation (eq. 14) - OLS / fixed effects

	Dependant variable: $\log(migr_{jit}/\sum_{i'\neq j} migr_{ji't})$							
	The Netherlands		The Netherlands		Italy- Total		Italy- North	
Total employment	0.652^a	–	0.584^a	–	0.936^a		0.945^a	–
log $L_{i(t-1)}$	(0.016)		(0.017)		(0.019)	–	(0.021)	
Service employ.		0.101		-0.013	–	0.537^a	–	0.326^a
log $L^y_{i(t-1)}$	–	(0.078)		(0.085)		(0.038)		(0.053)
Indus. employ.		0.506^a	–	0.558^a	–	0.254^a	–	0.681^a
log $L^x_{i(t-1)}$	–	(0.072)		(0.078)		(0.035)		(0.047)
Agri. employ.		0.144^a	–	0.152^a	–	0.162^a	–	-0.168^a
log $L^z_{i(t-1)}$	–	(0.041)		(0.045)		(0.017)		(0.047)
Prob-wage	$-0.462^{a\S}$	$-4.489^{a\S}$	–	–	-0.407^a	0.183	0.801^a	0.705^a
log $w_{i(t-1)}E_{i(t-1)}$	(0.414)	(0.411)			(0.089)	(0.121)	(0.136)	(0.136)
Distance	-1.050^a	$-1,109^a$	-1.146^a	-1.216^a	-0.231^a	-0.241^a	-0.836^a	-0.883^a
$\log(d_{ij})$	(0.035)	(0.038)	(0.037)	(0.040)	(0.024)	(0.025)	(0.036)	(0.040)
No-border	-0.386^a	-0.353^a	-0.353^a	-0.313^a	-0.809^a	-0.789^a	-0.522^a	-0.508^a
F_{ij}	(0.034)	(0.035)	(0.037)	(0.038)	(0.034)	(0.035)	(0.038)	(0.038)
Surface	0.051	-0.149^a	0.137^a	-0.095	-0.024	-0.062	-0.156^a	-0.081^b
$\log(S_i)$	(0.035)	(0.057)	(0.037)	(0.062)	(0.035)	(0.039)	(0.035)	(0.040)
a_j (sign)	(+)	(+)	(-)	(-)	(-)	(-)	(-)	(-)
Nb. Obs.	630	630	630	630	3366	3366	2033	2033
R2	0.9909	0.9912	0.9891	0.8912	0.9782	0.9782	0.9841	0.9847
F-Stat	3948	3631	3486	3631	6244	5727	5231	5036
Schwarz Criterion	290.7	284.4	400.8	390.4	5856.9	5891.2	3014.8	2946.0

Standards errors in parenthesis; ^{a}, ^{b}= significance at 1 and 5% levels

§: significant multicollinearity affects the fixed effect and the coefficients relating to Prob-wage.

Table 4: Non-linear Least Squares / Fixed Effects

	Dependant variable: $\log(migr_{jit}/\sum_{i' \neq j} migr_{ji't})$							
	Germany		Germany		Spain		Great Britain	
	model 1	model 2	model 1	model 2	model 1	model 2	model 1	model 2
σ_x	$3.740^{\#}$	$5.130^{\#}$	$3.850^{\#}$	$5.250^{\#}$	$1.534^{\#}$	$1.601^{\#}$	$1.304^{\#}$	$1.534^{\#}$
(elast. of substitution)	(0.663)	(0.999)	(0.674)	(1.001)	(0.159)	(0.216)	(0.108)	(0.229)
δ	3.619^a	2.490^a	3.760^a	2.560^a	0.461^a	0.621^a	1.536^a	1.018^a
(transport cost)	(0.524)	(0.351)	(0.521)	(0.348)	(0.161)	(0.164)	(0.210)	(0.138)
α_1	0.723^a	0.716^a	0.722^a	0.715^a	0.904^a	0.909*	0.725^a	0.724^a
(service employment)	(0.031)	(0.031)	(0.031)	(0.031)	(0.012)	(0.012)	(0.022)	(0.022)
α_2	-0.079	-0.057	-	-	-0.394^a	-0.323^a	0.205^a	0.208^a
(prob. wage)	(0.073)	(0.073)			(0.046)	(0.044)	(0.023)	(0.023)
λ	0.922^a	0.922^a	0.923^a	0.923^a	0.764^a	0.752^a	0.475^a	0.475^a
(migration cost)	(0.040)	(0.040)	(0.04)	(0.04)	(0.035)	(0.035)	(0.036)	(0.036)
b	0.862^a	0.856^a	0.851^a	0.848^a	1.412^a	1.453^a	1.265^a	1.267^a
(no-border)	(0.145)	(0.145)	(0.143)	(0.144)	(0.220)	(0.229)	(0.289)	(0.288)
α_4	0.684^a	0.700^a	0.703^a	0.715^a	0.036^a	0.037^a	0.235^a	0.235^a
(surface)	(0.079)	(0.079)	(0.078)	(0.079)	(0.014)	(0.015)	(0.041)	(0.041)
Nb. Obs	900	900	900	900	2310	2310	540	540
R^2	0.8798	0.8786	0.8796	0.8785	0.8271	0.8269	0.8806	0.8817
MSE	0.1991	0.2011	0.1992	0.2010	0.2472	0.2475	0.0720	0.0713
Schwarz Criterion	1211.6	1220.5	1206.2	1214.4	3495.6	3499.6	212.1	206.8

White consistent standard errors in parenthesis; a = significance at 1% levels; $^{\#}$= greater than 1 at 1% level.

Table 5: Non-linear Least Squares / Fixed Effects

	Dependant variable: $\log(migr_{jit}/\sum_{i'\neq j} migr_{ji't})$							
	The Netherlands		The Netherlands		Italy - Total		Italy - North	
	model 1	model 2	model 1	model 2	model 1	model 2	model 1	model 2
σ_x	$4.316^{\#}$	$5.630^{\#}$	$3.647^{\#}$	$4.626^{\#}$	$3.579^{\#}$	$4.165^{\#}$	$2.751^{\#}$	$3.283^{\#}$
(elast. of substitution)	(1.101)	(1.715)	(0.709)	(1.060)	(0.381)	(0.430)	(0.326)	(0.392)
δ	1.416^a	1.074^a	1.783^a	1.366^a	3.545^a	2.357^a	2.711^a	1.915^a
(transport cost)	(0.298)	(0.241)	(0.299)	(0.241)	(0.160)	(0.106)	(0.226)	(0.153)
α_1	0.463^a	0.469^a	0.429^a	0.435^a	0.974^a	0.963^a	0.936^a	0.945^a
(service employ.)	(0.027)	(0.029)	(0.027)	(0.027)	(0.020)	(0.019)	(0.022)	(0.022)
α_2	-0.454^a	-0.453^a	–	–	-0.059^b	0.043	0.044	0.073^b
(prob. wage)	(0.077)	(0.077)			(0.027)	(0.027)	(0.036)	(0.036)
λ	1.019^a	1.023^a	1.022^a	1.027^a	0.313^a	0.291^a	0.819^a	0.812^a
(migration cost)	(0.041)	(0.041)	(0.042)	(0.041)	(0.023)	(0.023)	(0.037)	(0.037)
b	0.514^a	0.508^a	0.516^a	0.509^a	9.044^a	11.836^a	0.884^a	0.901^a
(no-border)	(0.075)	(0.074)	(0.077)	(0.076)	(2.632)	(3.77)	(0.129)	(0.132)
α_4	0.469^a	0.452^a	0.601^a	0.584^a	0.032	0.047	0.001	-0.019
(surface)	(0.078)	(0.079)	(0.077)	(0.078)	(0.034)	(0.032)	(0.035)	(0.034)
Nb. Obs	630	630	630	630	3366	3366	2057	2057
R^2	0.9135	0.9136	0.9086	0.9088	0.7851	0.7932	0.8599	0.8591
MSE	0.0875	0.0873	0.0922	0.0921	0.3489	0.3357	0.2366	0.2379
Schwarz Criterion	356.1	355.2	384.1	383.2	6073.1	5943.1	3044.4	3056.4

White consistent standard errors in parenthesis; ^{a}, ^{b}= significance at 1 and 5% levels; $^{\#}$= greater than 1 at 1% level.

Table 6: Breaking points (Threshold relative distances)

	Germany	Spain	Italy	The Netherlands	Great Britain
Model 1	1.21	BH	1.63	1.56	BH
Model 2	1.35	BH	2.49	2.00	BH

BH = *black hole condition*

Figure 1: Rates of emigration: $Migr_{ji}/L_j$. ($w_i = w_j$; $d_{ii} = d_{ij} = 100$)