Baláž, Vladimír; Karasová, Katarína

**Article**

# Geographical gatterns in the intra-European migration before and after Eastern enlargement : the connectivity approach

Ekonomický časopis

**Provided in Cooperation with:**
Slovak Academy of Sciences, Bratislava

*Reference:* Baláž, Vladimír/Karasová, Katarína (2017). Geographical gatterns in the intra-European migration before and after Eastern enlargement : the connectivity approach. In: Ekonomický časopis 65 (1), S. 3 - 30.

# Geographical Patterns in the Intra-European Migration before and after Eastern Enlargement: The Connectivity Approach[1]

*Vladimír BALÁŽ – Katarína KARASOVÁ**

**Abstract**

*This paper analyses spatial patterns in intra-European migration flows in the periods 1997 – 2004 and 2005 – 2013. The paper uses network analysis, and regression and factor analysis in order to establish the major determinants of the spatial patterns exhibited by intra-European migrant stocks. The EU's Eastern enlargement generated vast East-West migrant flows and prompted a particular reconfiguration of the migration network. The basic topology of the network, however, did not change across the two observed periods: The whole network remains dominated by a 'rich club' structure. The topology of the network was seen to rely on a complex and stable set of long-term institutions, such as culture and language structures, and/or established pathways of trade in goods and knowledge.*

**Keywords:** *intra-European migration, network analysis, connectivity models, push-pull models, the EU Eastern enlargement*

**JEL Classification:** F22, F55, D85

## 1. The Intra-European Migration System

### 1.1. Conceptualisation

The conventional models concerning international migration originate in the human capital theory and focus on decisions by individual migrants (Sjaastad, 1962; Harris and Todaro, 1970; Borjas, 1987); the human capital approach considers international migration as an investment decision. The present discounted

* Vladimír BALÁŽ – Katarína KARASOVÁ, Institute for Forecasting, Slovak Academy of Sciences, Šancová 56, 813 64 Bratislava 1, Slovak Republic; e-mail: vbalaz@yahoo.com; katka.karasova@centrum.sk

[1] This project has received funding from the European Union's Horizon 2020 under Grant Agreement No. 649491 – YMOBILITY.

value of expected returns on a migrant's human capital abroad must exceed the expected returns in the country of the migrant's origin. Income and employment opportunities are considered to be the major drivers of voluntary trans-border movements. The classical theories of international migration recognise that there are some non-monetary ('psychological') costs and returns associated with international migration, such as opportunity costs of time, and/or to loosing proximity to familiar surroundings, family, and friends. The 'psychological' costs are 'surely significant, probably far more so than the money costs' (Sjaastad, 1962, p. 84). The 'psychological costs and returns', however, are almost never quantified in migration research (but see De Jong, Chamratrithirong and Tran, 2002; Olgiati, Calvo and Berkman, 2013 for notable exceptions).

The 'new economics of labour migration' (Mincer, 1978) considers international migration a risk-sharing and resource-diversification strategy which is undertaken by migrants' families. The novel economics of the labour migration and welfare approach assumes that migrants base their choice of destination country on an effective cost/benefit analysis, whereby economic variables play a major role in individual and family decisions made on migration. Borjas (1999), for example, suggested welfare payments as a pull factor for potential immigrants. The research on 19 European countries over the period 1993 – 2008, however, indicated little support for the existence of a welfare magnet effect on the inflows of non-EU immigrants (Giulietti et al., 2013).

The cost/benefit approach also has some limitations. The focus on the economic motives related to migration operates best when explaining migration flows from less developed countries to developed ones. The cost/benefit approach may be less powerful for explaining flows between the high-income and middle-high income countries. The intra-European migration system, for example, encompasses a much more diverse set of migration motives than simply those of job and income disparities. Some recent empirical studies on the causes of intra-European migration indicate that migration within Europe is motivated primarily by work--related and family-related factors, but on the whole various family-related and social motivations are mentioned more frequently than solely work-related reasons (Verwiebe, Wiesböck and Teitzer, 2014, p. 129). As noted by Castro-Martín and Cortina (2015, p. 110) '*Across Europe we find mobility taking place at different times in people's lives: one can find students, working adults and pensioners all in the process of migrating.*' Thus decisions concerning intra-European movements are more complex and may involve more than the traditional economic motives (King, 2002; Favell, 2008). Contemporary intra-European migrants are very different from those of the 1960s and 1970s. Current migrants tend to be younger and better educated than those of the classical period of intra-European labour

migration. Migration between high-income countries and middle high-income may reflect quite varied motivations/tastes and lifestyle choices, such as education, language studies (Parutis, 2014), novelty seeking, personal relationships, cultural preferences, climate considerations and many more. The same pair of European countries may therefore generate quite diverse forms of migrant exchange. Flows of Portuguese labour migrants pursuing higher wages and Portuguese students enrolling in British universities, for example, meet flows of UK retirees seeking the sunnier climate and lower living cost of living in Portugal.

This paper adopts the concept of the world migration system (Zlotnik, 1999). The world system theory assumes that the geographical layout of the migrant stocks reflects centre/periphery relations in the world system. The world migration system consists of a group of receiving (usually economic-core) countries that are linked to a set of sending countries by relatively large flows and counter-flows of migrants. We argue in this paper that the system of the intra-European network is a specific substructure of the world migration system (DeWaard, Kim and Raymer, 2012). The most tangible benefit of the system approach is that it does not consider migration flows as mere reactions to income and employment differentials, but instead it 'forces the researchers to consider both origin and destination contexts and the relationship between them' (Bakewell, 2014). The concept of a migration system highlights 'the diverse linkages between places, including flows of information, goods, services and ideas, as well as people' (Fawcett, 1989, p. 673). The geographical layout of migration flows is determined by a set of political, economic, linguistic and cultural relations between periphery and centre (the latter often is a former colonial power). The basic structure of the institutions governing migration flows is relatively stable, but these institutions are subject to dynamic change. Institutional change may either reduce or reinforce migration flows within the migration system. Increases in flows between two countries, for example, may lead to the development of more frequent and cheaper travel connections. Improvements in travel, in turn, reinforce migration flows and increase migrant stocks. Increases in migrant stocks may strengthen cultural linkages between countries and lower, for instance, the psychological costs of migration (Jenissen, 2007, p. 433).

We consider the original contribution of this paper to be in identifying the importance of the connectivity factors in determining the geographical layout of the intra-European migration network. Our research indicates that connectivity factors are more important than push-pull factors as determinants of the geographical distribution of the intra-European migrant stocks; this paper is methodologically innovative in its analysis of the role of connectivity factors. While this is not the first work to highlight the importance of connectivity factors (see

for example Fawcett, 1989; Jenissen, 2007) our analysis is far more strongly focussed on connectivity, and methodologically incorporates more variables (see Table A1) than many other macros-scale analyses.

The first stage of the analysis examines general trends in spatial polarisation, and how these have changed over time, using both descriptive statistics and network theory analysis. The network analysis brings a fresh perspective to the spatial analysis. To the best of our knowledge this is the first time that network science has been used to analyse the intra-European migration network system. Three different metrics are used to analyse the network topology. The modularity and density statistics indicate that while the overall density of the intra-European migration network has increased, the network has also become more fractured into particular large-scale clusters (modules). The basic topology of the intra--European migration network, however, did not change much across the periods 1997 – 2004 and 2005 – 2013.

The second stage uses traditional methods of regression and factor analysis to establish the major determinants of the geographical layout of the intra-European migrant stocks. Most quantitative studies on national and regional levels refer to the conventional push-pull variables such as unemployment and economic activity rates, GDP disparities, household incomes and education (Pedersen, Pytliková and Smith, 2008; Fagiolo and Mastrorillo, 2013; Landesmann and Leitner, 2015). The network effects usually are modelled via language dummies, former colonial relationships (Hooghe et al., 2008) and/or spatial distances (Dennett and Wilson, 2013; Sardadvar and Rocha-Akis, 2016). In this study, we try to go beyond the conventional push-pull framework. We recognise the diversity of migration motivations and explore the potential of novel data sources for explaining the geographical layout of the intra-European migration network. The set of conventional push-pull variables is completed with data on the psychological and social determinants of migration (coming from the European Social Survey), and data on network structures and connectivity in terms of geographic, economic, and linguistic distance. A focus on the intra-European migrant stocks helped to overcome some limitations conventional gravity models encounter when modelling migration from less developed countries to developed ones. The Eurostat, OECD, European Social Survey and Eurobarometer surveys provide a rich source of data on the monetary, non-monetary and network determinants of migration between the 31 European countries.[2]

[2] The 31 countries include Austria (AT), Belgium (BE), Bulgaria (BG), Croatia (HR), Cyprus (CY), Czech Republic (CZ), Denmark (DK), Estonia (EE), Finland (FI), France (FR), Germany (DE), Greece (GR), Hungary (HU), Ireland (IE), Iceland (IS), Italy (IT), Luxembourg (LU), Netherlands (NL), Latvia (LV), Lithuania (LT), Malta (MT), Norway (NO), Poland (PL), Portugal (PT), Romania (RO), Slovakia (SK), Slovenia (SI), Spain (ES), Sweden (SE), Switzerland (CH), and United Kingdom (UK).

We propose three hypotheses:

H1: *Prior to the Eastern enlargement, income and employment differentials were relatively low within the EU and insignificant in terms of explaining the geographical patterns of intra-European migrant stocks. The income and employment differentials became significant after the EU's eastern enlargement and the opening of access to labour markets by the EU-15 Members*.

H2: *The intra-European migrant stocks are responsive to connectivity variables such as established patterns of trade in goods, knowledge exchange and tourist flows. The connectivity patterns reflect the centre-periphery economic and cultural relations between key European powers (UK, FR and DE) and have remained relatively stable over decades.*

H3: *Rapid and significant increases in the total migrant stocks of selected destinations, inter alia, are explained by the increased importance of language considerations. Language competences were important both in relation to neighbour-country flows and the UK-centred flows. Increased levels of migrant stocks in the UK are associated with the establishment of English as a global language and also with the dominant position British universities have in the European market for tertiary education.*

### 1.2. Geography of the Intra-European Migration System

There has been significant growth in the stocks of intra-European migrants in the last two decades. The average annual stock increased from 9.1 million in the period 1997 – 2004 to 13.7 million in the period 2005 – 2013 in the EU (Eurostat, 2016). The spatial distribution of the intra-European migrant stocks is quite uneven and follows an exponential distribution. The average stocks were 10,882 and 16,085, but the median ones were 833 and 854 in the periods 1997 – 2004 and 2005 – 2013 respectively. As was observed, the 2004 Eastern enlargement generated substantial increases in total stocks, but the stock distributions changed little.

The Eurostat and the OECD data indicate that the intra-European migrant stocks are highly concentrated. Six migration destinations (UK, FR, DE, CH, IT, ES) attracted over 75.4% and 75.9% of all intra-European migrants in the periods 1997 – 2004 and 2005 – 2013 respectively. Fifteen destinations attracted 95.9% and 95.8% of all intra-European migrants in the periods 1997 – 2004 and 2005 – 2013 respectively.[3] We further refer to the abovementioned six countries as the 'core' and fifteen countries as the 'centre'. Shares of the core in total migrant stocks were broadly stable throughout the period 1997 – 2013, despite the overall growth in numbers and the EU's eastern enlargement, suggesting the persistence of some underlying determinants.

[3] The centre is defined to be the AT, BE, DE, DK, FI, FR, IE, IT, LU, NT, NO, ES, SE, SW and UK.

The overall volume of the intra-European stock of migrants increased 1.52 times across the periods 1974 – 2004 and 2005 – 2013. Table 1 highlights three main trends in the geographical structure of the intra-European migrant stocks:

- First, centre to centre stocks accounted for the largest share of total stocks in the 1997 – 2004 period – by far. Following the consecutive waves of EU enlargements in the 2000s that effectively incorporated new countries into the EU's labour market, the numbers of migrants from these new countries increased sharply between 1997 – 2004 and 2005 – 2013: 2.87 times. Consequently, the volume of periphery-centre stock movements almost matched centre-centre movements in the period 2005 – 2013. Centre to periphery and periphery-periphery movements were quite low, and generated about 4.2% of total stocks in both periods.
- Secondly, the 2004 EU Eastern enlargement generated substantial increases in migrant stocks after 2004. Stocks originating in the East accounted for 20.1% of the total stocks in the period 1997 – 2004, but 38.0% in the period 2005 – 2013. Eastern originating stocks partially replaced those originating in the South. The share represented by the South-originating stocks dropped from 39.4% in 1997 – 2004 to 26.6% in 2005 – 2013. The Eastern enlargement of the EU was the most significant event affecting intra-European migration – since the 1960s. Free movement of workers helped to ease the imbalances which existed across the EU labour markets (Kahanec, Pytliková and Zimmermann, 2016; Kahanec and Zimmermann, 2016) and was reflected in substantial increases in the numbers of EU-10 nationals in the EU-15 countries. Some EU-15 Members opted for gradual access to their labour markets by EU-10 citizens.[4] The gradual opening of labour markets was reflected in uneven increases in the numbers of EU-10 nationals in the EU-15 countries. The average annual stocks of EU-10 nationals in the EU-15 countries increased from 1.82 to 5.23 million across the periods 1997 – 2004 and 2005 – 2013 (Table 1). Income-driven migration, no doubt, generated substantial stocks of immigrants in the EU-15 countries, but the increase in migrant stocks had some distinctive geographical patterns. The UK, Italy and Spain, for example, received by far the highest inflows of immigrants from the East. This may indicate the importance of language considerations in the geographical layout of intra-European migrant flows. The impact of cross-border flows, on the other hand, was less important than expected (see also Galgóczy, Leschke and Watt, 2009, p. 14).

[4] The EU-10 countries include the Czech Republic, Slovakia, Poland, Hungary, Slovenia Lithuania, Latvia, Estonia, Romania and Bulgaria. Ireland, Sweden and the UK opened their labour markets immediately after EU eastern enlargement in 2004. Some EU-15 countries negotiated gradual adaptation of their labour markets: Greece, Spain, Italy, Portugal and Finland opened in 2006, Luxembourg in 2007, France in 2008, and Belgium and Denmark in 2009 and Germany and Austria in 2011. Romania and Bulgaria obtained access to labour markets in Italy in 2012 and in the UK in 2014.

• Thirdly, intra-European migrant exchange was informed also by geographical and linguistic proximity. Neighbour-country exchanges accounted for some 40.2% and 31.9% of the total exchanges in the periods 1997 – 2004 and 2005 – 2013 respectively. The decrease in the relative importance of neighbour-country exchanges between the two periods is related to the significance of long-haul migration from the Eastern members of the EU to the UK, Italy, and Spain after 2004. Knowledge of the host-country language is an extremely important precondition for finding a job and/or integration into the host country's society. Here we apply two concepts of linguistic proximity. The narrow concept refers to pairs of origin-host countries where at least 10% of each population speaks the same language (e.g. Germany and Switzerland). The broad concept applies to pairs of origin-destination countries where at least 10% of each population speak a language from the same language family (Germanic, Romance, Slavic, Greek or Finno-Ugric). We assume, here, that it is easier to understand and learn a language from the same language family than it is to learn/understand a language from a different language family. The intra-European migrant stocks which had a broad language proximity to their host country generated over 50% of total migrant stocks in both periods, which indicates the importance of language considerations in migration decisions.

T a b l e 1

**Average Annual Intra-European Migrant Stocks** (million persons and per cent of total)

| **Flow type** | **1997 – 2004** | **2005 – 2013** | **Growth rates: 2005 – 2013 to 1997 – 2004** |
|---|---|---|---|
| Total stocks: | 9,064,398 | 13,736,256 | 1.52 |
| Stocks by position within the migration system | | | |
| centre to centre | 5.57 (61.5%) | 6.64 (48.4%) | 1.19 |
| centre to periphery | 0.13 (1.5%) | 0.22 (1.6%) | 1.68 |
| periphery to centre | 3.12 (34.4%) | 6.52 (47.5%) | 2.09 |
| periphery to periphery | 0.24 (2.7%) | 0.35 (2.6%) | 1.44 |
| Stocks by region of origin: | | | |
| Middle Europe | 2.55 (28.2%) | 3.33 (24.2%) | 1.30 |
| Eastern Europe | 1.82 (20.1%) | 5.23 (38.0%) | 2.87 |
| Northern Europe | 1.12 (12.3%) | 1.53 (11.1%) | 1.37 |
| Southern Europe | 3.57 (39.6%) | 3.65 (26.6%) | 1.02 |
| Stocks by geographical and language proximity | | | |
| neighbour countries | 3.64 (40.2%) | 4.38 (31.9%) | 1.20 |
| language proximity (narrow\| | 1.96 (21.6%) | 2.21 (16.1%) | 1.13 |
| language proximity (broad) | 4.88 (53.8%) | 6.94 (50.5%) | 1.42 |

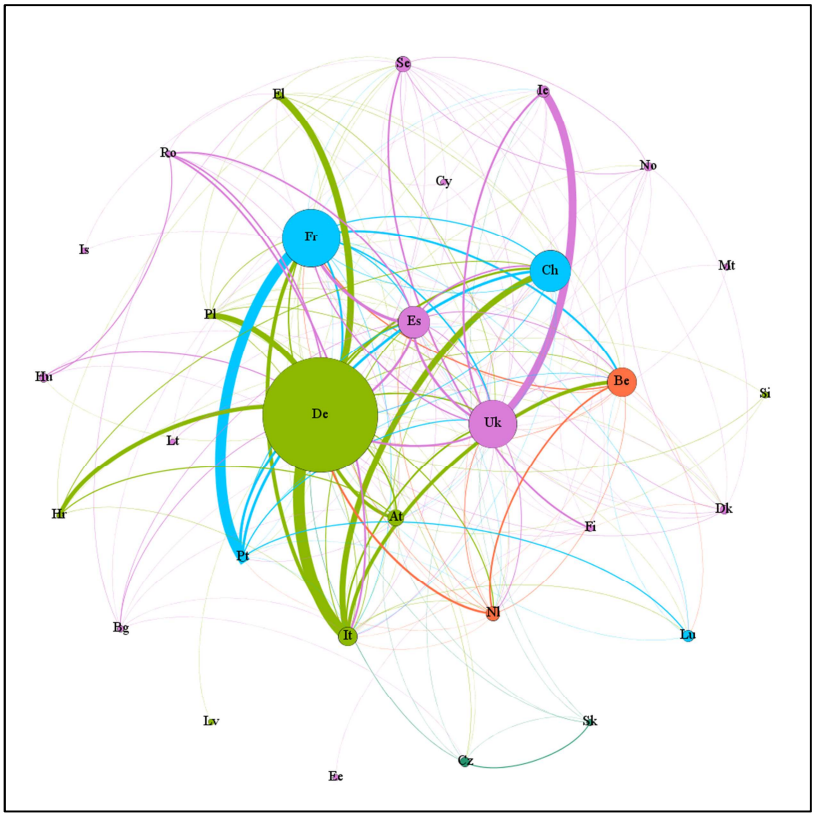
*Notes*: The centre is defined to be the AT, BE, DE, DK, FI, FR, IE, IT, LU, NT, NO, ES, SE, SW and UK. All other countries are considered periphery countries of the intra-European migration network. East is defined as CZ, HU, PL, SK, SI, LT, EE, LV, RO and BG. South is defined as ES, IT, PT, EL and CY. North is defined as UK, IS, DK, NO, SE, and FI. Middle is defined as DE, FR, BE, NL, LU, CH and AT. Countries separated by sea distance were considered neighbours if connected via bridge (DK-SE) or tunnel (UK-FR) and when the sea distance was shorter than 100 km. Language proximity (narrow) was established for countries where at least 10% of each population spoke the same language. The broad concept applies to pairs of origin-host countries where at least 10 % of the population of each country spoke a language from the same language family (Germanic, Romance, Slavic, Greek and Finno-Ugric).

*Source*: Authors' computation based on the Eurostat and OECD data for 31countries.

The overall pattern of centre and periphery exhibited both continuity and change in 1997 – 2013. The continuity is evident in the persistent importance of the six destination countries, despite the overall rapid expansion of migration flows. However, there were also changes in the importance of individual centre countries, and in their connectivity to their peripheries. Italy, Spain and the UK were the major targets for intra-European migration. Italy and Spain emerged as principle destinations for inflows from Romania after 2004 (Figures 1 and 2). The UK accounted for a substantial increase in the absolute stocks of intra-European migrants (1.95 times). In contrast, stocks increased by only 1.07 times in France, and 1.05 times in Germany between 1997 – 2004 and 2005 – 2013. This development documents the increasing importance of English as a global language and the dominant position of the UK on the European market for tertiary education (Hypothesis 3).

F i g u r e 1

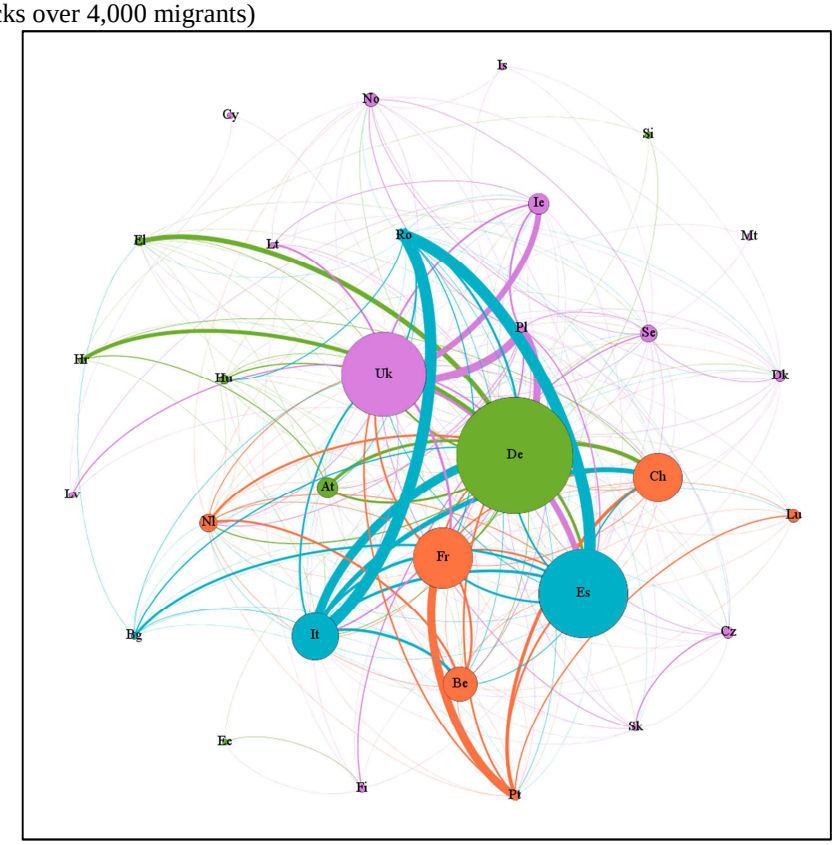
**Network Diagram for Intra-European Migrant Stocks in 1997 – 2004**
(stocks over 4,000 migrants)




*Note:* All flows have the colours of their sending modules.

*Source*: Authors' computations.

F i g u r e 2
**Network Diagram for Intra-European Migrant Stocks in 2005 – 2013**
(stocks over 4,000 migrants)




*Note:* All flows have the colours of their sending modules.
*Source*: Authors' computations.

## 2. The Network Analysis

Focussing on the changes over time, the migrant stocks are analysed using a network science approach (Easley and Kleinberg, 2010), utilising the concepts of nodes, edges, degrees and weighted degrees. These considerations were applied to the matrix of inflows and outflows from the 31 European countries.

This chapter first examines the topology of the intra-European migrant networks – the construction of the networks and how the different nodes in each network are connected to each other and how they communicate. The nodes of the networks are European countries sending/receiving stocks of European citizens. Each node has a degree, which is the number of connections (edges) the node has to other nodes. Countries with a high degree of centrality are more interconnected within the intra-European migration exchange. In directed networks

(such as flows of migrant between countries), each node (country) has two degrees: an in-degree and an out-degree. Any outgoing flow from country *A* is an ingoing flow to country *B, C, D,…,N*, so that the mean in-degree and the mean out-degree are always equal in directed graphs. The degrees can be weighted by edge thickness (volumes of flows). The weighted in-degree of a node *A*, for example, is the sum of all stocks of intra-European migrants in country *A*.

Three different metrics were used to analyse the network topology. The different metrics offer different views on the geographical pattern of intra-European migration:

- Metric **A** considers only stocks higher than 4,000 migrants and so essentially maps the largest migrant stocks in Europe. There were some 201 stocks with 4,000+ migrants in 1997 – 2004.[5] These 201 stocks accounted for 95.7% of the total migrant stocks in the abovementioned period. In 2005 – 2013 the number of stocks with 4,000+ migrants increased to 255. These 255 stocks accounted for 96.7% of the total migrant stocks in 2005 – 2013. Metric A was used to produce Figures 1 and 2. These diagrams of stocks over 4,000 only are presented, in particular, for reasons of clarity. The full depiction of the network would require 31 x 30 = 930 edges, most of which would be quite small. As for the quantitative analysis, metric A favours big countries with high number of sizeable stocks, but works less for small countries.
- Metric **B** considers stocks of over 0.1% of the sending country's population and measures the intensity of outmigration from each sending country. Some small countries generate few flows with 4,000+ out-migrants, yet they have a significant intensity of outmigration. There were some 129 migrant stocks representing over 0.1% of the sending country's population in 1997 – 2004 but 176 in 2005 – 2013. Metric B favours small countries with relatively high outflows of migrants, but works less well for big countries.
- Metric **C** records the five largest stocks by the sending countries and describes the geographical preferences of the emigres. Metric C is a compromise between metrics A and B. Total 155 stocks in metric C encompass 85% of all intra-European stocks. The metric C is used in the quantitative analysis in chapter 3. The metric C copes with lower data limitations (related to explanatory variables) than metrics A and B.

Statistics for the metrics A, B, and C are presented in Table 2. There were four main findings.

First, there is a very high level of interconnection within the 31 countries. Starting at the simplest level of analysis, there was an increase in the *average degree* (average number of connections for each country) from 12.968 in 1997 –

[5] Threshold 4,000+ was chosen as to capture at least 95% of the total flows in each time period.

2004 to 16.452 in 2005 – 2013 for metric A, and from 8.323 to 11.355 for metric B. This trend was also reflected in an increase in *graph density:* that is, in the proportion of the maximum possible number of edges from a node (on a scale of 0 to 1). The graph density increased from 0.216 to 0.274 between 1997 – 2004 and 2005 – 2013 for metric A and from 0.139 to 0.189 for metric B. An increase in the average degree was accompanied by an increase in the *average weighted degree* (average stocks of intra-European migrants in each country) for both metrics A and B. The network density, of course, remains constant if only the five largest stocks in terms of sending countries are considered (metric C). The modularity and average clustering coefficients, however, increased also for metric C despite the constant density of the network.

Secondly, confirming the evidence in Table 1, there was an increase in the values of the average clustering coefficient which indicates how nodes are embedded in their neighbourhood. The average clustering coefficient (scale 0 to 1) increased between the two periods for all three metrics. The statistics on weighted degrees, network density and clustering indicate that the intra-European migrant exchange (i) increased in absolute terms, (ii) diversified in terms of the numbers of countries and flows involved in the exchange, and (iii) improved in terms of the interconnectivity among European countries.

T a b l e 2

**Network Topology for Intra-European Migrant Stocks in 1997 – 2013**

| **Network overview** | **1997 – 2004** | **2005 – 2013** |
|---|---|---|
| **Metric A: Stocks over 4,000 persons from sending country** | | |
| No. of observed stocks | 201 | 255 |
| Average degree | 12.968 | 16.452 |
| Average weighted degree | 279,643 | 428,564 |
| Graph density | 0.216 | 0.274 |
| Modularity | 0.260 | 0.257 |
| Avg. clustering coefficient | 0.561 | 0.571 |
| **Metric B: Stocks over 0.1% of the sending country's population** | | |
| No. of observed stocks | 129 | 176 |
| Average degree | 8.323 | 11.355 |
| Average weighted degree | 234,918 | 391,328 |
| Graph density | 0.139 | 0.189 |
| Modularity | 0.394 | 0.419 |
| Avg. clustering coefficient | 0.401 | 0.457 |
| **Metric C: Five largest stocks from sending country** | | |
| No. of observed stocks | 155 | 155 |
| Average degree | 5 | 5 |
| Average weighted degree | 251,364 | 370,399 |
| Graph density | 0.167 | 0.167 |
| Modularity | 0.273 | 0.303 |
| Avg. clustering coefficient | 0.480 | 0.520 |

*Source*: Authors' computations using Gephi software.

The third finding relates to the structure of the network. *Modularity statistics* have been used to measure the strength of the division of the network into communities (modules or clusters).[6]

Communities are characterised by dense connections between nodes within the communities, and sparse connections with other nodes. The modularity of the network slightly decreased for metric A and slightly increased for metrics B and C between the two observed periods. The numbers of modules dropped from 5 to 4 for metric A between the two periods. Metric A is the most relevant for the modularity statistics, as it covers over 95% of the total intra-European migrant stocks.

Fourthly, in some complex networks (e.g. international trade networks), the high-degree nodes tend to form 'cliques' which have more interactions with equal or higher degree nodes – the so-called 'rich-club' phenomenon (Barrat et al., 2004, p. 3751).

The rich club concept refers to the confining of exchanges to those between small numbers of big players. Fagiolo and Mastrorillo (2013) analysed the United Nations database of more than 230 destination countries and territories from across the globe in the period 1960 – 2000. They found the world-wide migrant network poorly concentrated with no evidence of a 'rich-club' phenomenon. Major destination countries receive most migrant inflows from a high number of developing countries. The contributing countries, in turn, have on average a low number of inflows and/or receive a small stock of migrants. The network of intra-European migrants is quite different from the world-wide network of migrants: a 'rich club' pattern is very evident. Exchanges between the six major destinations accounted for 61.5% and 48.4% of the total flows in the two observed periods (Table 1). The modularity and density statistics indicate that while the overall density of the intra-European migrant network increased across these periods, the network became more fractured into particular large-scale clusters.

The basic topology of the intra-European migrant network did not change much between the periods 1997 – 2004 and 2005 – 2013. The average migrant stocks per country accounted for fairly similar exponential distributions ($85{,}667^{e-0.011x}$ versus $58{,}933e^{-0.011x}$) and the whole network was dominated by a 'rich club' structure. The topology of the network relies on a complex and stable set of long-term institutions, such as culture and language proximity structures, and/or established pathways for trade in goods and knowledge (Hypothesis 2).

[6] The modularity measure is computed as the number of links in each community minus the number of links in the same groups in a graph where the links were redistributed randomly (Newman, 2006).

The major patterns exhibited by intra-European migrant stocks in the two time periods are visualized using network diagrams. The network diagrams produced for metric A help to identify the gravitation centres of migrant inflows and important communities (clusters). A layout algorithm places connected nodes closer together than unconnected ones and highlights clustering-by-attribute. The Fruchterman-Reingold (FR) layout algorithm was used to arrange the nodes and edges in a web (Figures 1 and 2).[7]

The network diagrams point to two main findings. First, they confirm visually, and in terms of individual edges and nodes, the broad patterns observed in Table 2.

There is a distinctive pattern of core and peripheries in both diagrams, where the core is formed by the UK, Germany, France, Switzerland, Italy and Spain in 1997 – 2013. Secondly, there are also strong centre-centre flows within the modules (DE-CH, DE-AT, FR-CH, BE-FR, IE-UK). Many these flows seem to be based on language proximity, geographical proximity and/or economic connectivity.

- Five modules emerged in the period 1997 – 2004: (1) the Germany-centred module with major inflows from Poland, Austria, Italy and Greece; (2) the UK and Spain-centred module – Ireland, the Nordic countries and Spain were major contributors to the UK, while Romania and Portugal were major contributors to Spain; (3) the France-centred module (with Portugal and Switzerland the main contributors); (4) the Belgium-Netherlands module; and (5) a small, but distinctive Czech-Slovak module (Figure 1).
- The Eastern enlargement initiated certain reconfigurations of the network modules. Four module configurations were established in the period 2005 – 2013: (1) The Germany-centred module changed relatively little – Austria, Hungary, Italy Greece and Poland remained the major contributors to stocks of foreign populations in Germany, while Germany increasingly contributed to Switzerland; (2) the UK-centred module accounted for a major increase in migrant stocks. In addition to the traditional contribution from Ireland, the UK benefited from influxes of Poles, Czechs, Slovaks and citizens of Baltic and Scandinavian countries; (3) Italy and Spain formed a distinctive southern-European module with major contributions from Romania and Bulgaria; and (4) The France-centred module merged with the Belgium-Netherlands module (Figure 2).

[7] The FR algorithm belongs to a family of force-directed layout algorithms. These algorithms use attractive and repulsive forces to minimize the energy of the system by moving the nodes and changing the forces between them. In the FR the sum of the force vectors determines the direction a node should move. The nodes' positions stabilise when the energy of the system is minimized and the system reaches an equilibrium state (Fruchterman and Reingold, 1991).

## 3. The Determinants of Intra-European Migrant Stocks

### 3.1. The Sample and Variables

The second stage involved the modelling of the determinants of the migration stocks distributed between the 31 European countries. The main data source in this paper is the Eurostat database on European citizens living in other (non-native to them) EU member states. Data on the dependent and independent variables were only available for 31 European countries for the period 1997 – 2013. Some small countries (including Lichtenstein, Macedonia, and Serbia) were excluded from the analysis because of data limitations.

Most studies on international migration use gravity models based on annual panel data, and consider either all or only the largest binary flows between countries of origin and destination (Pedersen, Pytliková and Smith, 2008; Mayda, 2010; DeWaard, Kim and Raymer, 2012). Gravity models based on panel data, however, have their own problems. They usually do not go beyond income differentials, employment opportunities, and a limited set of set of other explanatory variables, such as population size and/or physical distance – for a cross-section of sending and receiving countries. The gravity models and pooled ordinary least squares (OLS) estimators used to neglect the individual heterogeneity amongst the different countries involving factors such as culture, language, geography etc. that might have an important impact on migration (Alvarez-Plata, Brücker and Siliverstovs, 2003, p. 20). A focus on the largest binary flows helps to explain a substantial part of international migration flows. However, it is of less use in explaining the geographical distribution of flows. It also tends to neglect outflows from minor countries. Some large-scale studies consider all available flows between countries of origin and destination. Pedersen, Pytliková and Smith (2008), for example, analysed immigration flows to, and immigration stocks in, 26 OECD countries from 129 countries of origin. Such studies are necessarily limited by the lack of available data for many explanatory variables.

Instead of traditional gravity models, this study refers to an approach developed by Alvarez-Plata, Brücker and Siliverstovs (2003). It models the migration function as

$$mst_{fh} = f(EV)$$

where

$mst_{fh}$ – the share of emigrants from country,
$h$ – residing in country $f$,
EV – a vector of explanatory variables.

The removal of visa and travel barriers and improvements in the availability of cheap air travel have significantly boosted intra-European migrant exchange

since the 1990s. Every European country has migrants who come from other European countries. A full analysis of all the flows would have resulted in a matrix of 31 origins by 30 destination countries by 17 years. However, data for some years were missing, and there were only minor flows between mutually distant smaller countries (e.g. Latvia and Greece) resulting in erratic annual changes. Small stocks of migrants from more distant countries sometimes are underreported, or represent missing data. This study therefore does not aim at analysing the full matrix of origin by destination countries. It opts for a rather different approach and computes the proportion of the total migrant stocks of the top five destination countries which come from each country of origin (metric C from chapter 2, and Table 2). The five major destinations for each country accounted for some 80%, and in some cases, for 90%, of the total outflows of each country. The final matrix comprised 31 sending countries by the five top destination countries, providing matrices of 155 cells for each time period. This set of matrices accounted for 85% of the intra-European migration found within the sample of 31 European countries – in 1997 – 2013. This approach is less powerful for explaining the total volume of migrant exchange, but works better for identifying the factors which are important for explaining the geographical distribution of migrant stocks. Annual averages were produced for two periods: 1997 – 2004 and 2005 – 2013. This also helped to overcome some of the limitations in the availability of time-series data for the independent variables, as discussed below.

The EU member states generate migration outflows of very diverse sizes and intensities. Germany, for example, generated some 1.06 million migrants who live in other European countries, but the proportion of out-migrants in the total German population was only 1.3% in the period 2005 – 2013. Latvia generated a mere 0.14 million emigrants, but the proportion of Latvian out-migrants in the total Latvian population was 6.4% for the same period. Some Eastern EU member states generated massive migration outflows both in absolute and relative terms. Romania, for example, had 1.92 million of its citizens, or 9.3% of its total population, abroad in 2005 – 2013.

We used two dependent variables when analysing the geographical distribution and intensity of migrant flows:

- Dependent variable 1 consists of the percentages of the total outmigration of a country going to each of the top five destination countries. This variable measures the simple geographical preferences of out-migrants. Dependent variable 1 proxies international migration flows by stocks of foreign nationals by country of birth, and therefore includes nationals who have already acquired destination-country citizenship. Where this data was unavailable, the ‘stock of foreign nationals by nationality’ was used instead.

• Dependent variable 2 represents the numbers of each origin country's citizens resident in the top five destination countries, weighted by the total population of the origin country. The variable therefore accounts for both geographical structure and the intensity of out-migration. This variable is also used by Alvarez-Plata, Brücker and Siliverstovs (2003).

This choice of explanatory variables reflected assumptions made according to the human capital and network theories of international migration. Pedersen, Pytliková and Smith (2008) found that migration flows between OECD member countries and developing countries responded to the traditional economic determinants (disparities in GDP and unemployment rates), but cultural and linguistic distance turned out to be important as well. Much more data are available for intra-European migration. In addition to the data on GDP, the Eurostat provides several measures of wages and unemployment, and also data on social benefits. We computed seven push-pull variables related to income motives for migration. These variables included four measures related to income gaps: (1) per capita GDP at purchasing power parity, (2) average net monthly wages for a single person with no children, (3) average net monthly wages for a married person with two children, and (4) average levels of social benefits. Variable (1) is a conventional measure which has been used to evaluate income disparities in many world-wide studies. Variables (2) and (3) can be used for direct measures of wage disparities. Variable (4) refers to Borjas's idea (1999) – that welfare payments are a pull factor for immigration from lower-income countries. Three push-pull measures targeted employment disparities between countries of migrant origin and destination: (5) total unemployment rates, (6) unemployment rate for young people up to age 25, and (7) the long-term unemployment rates.

A further group of push-pull factors is related to the monetary and non-monetary costs and returns of migration, drawing particularly on human capital theory (Sjasstad, 1962). Non-monetary factors influencing international migration are approximated by six variables drawn from rounds of the European Social Survey (ESS, 2002 – 2012). These are national averages for: life satisfaction; satisfaction with the quality of education; satisfaction with the current economic performance of a country; opinions on the state of the country's democracy; and self-reported levels of personal trust and happiness (variables 8 – 13). Each push-pull variable is expressed as a ratio of sending/destination country values for each particular indicator.

We further define nine connectivity variables. In this paper 'connectivity' is understood in terms of geographic, linguistic, and economic distance. Connectivity refers to the various communication channels which two or more countries use for the exchange of people, goods and knowledge (Jenissen, 2007). As for

economic connectivity, shares of total merchandise imports and exports provide a proxy for geographical patterns in international trade (variables 14 and 15). Trade and migration seem to be very interconnected. A meta-analysis of 48 studies on international trade and migration (Genc et al., 2011) indicated that immigration boosts trade, but the question remains as to what extent the estimated correlation is indicative of a truly causal effect. Anderson (2015) studied bilateral labour migration and trade flows in the EU-28 area. Results suggested that not only does migration affect trade, but also bilateral trade affects labour migration. Three indicators of patent flows were used as measures of knowledge connectivity: foreign ownership of domestic inventions – which is equivalent to the export of patents (variable 16); domestic ownership of foreign inventions – which is equivalent to patent imports (variable 17); and the numbers of patents having foreign co-inventors – a measure of mutual research co-operation (variable 18). Data on patents proxy flows of knowledge. We assume that there is an affinity between human migration flows and, for example, international scientific and technological co-operation and the knowledge spill-overs in business and research.

Personal (tacit) knowledge of a foreign country's people is approximated by tourism data: specifically the numbers of nights spent by nationals from destination countries in the migrants' home countries (variable 19). Two variables were used in relation to language knowledge. The 2005 and 2012 Eurobarometer surveys provided data on languages known, and languages considered useful to know other than mother tongues (variables 20 and 21). Spatial proximity (variable 22) was used as an indirect measure of transport costs, and was expressed as road distance (in km) between the capital cities of countries.

The importance of the push-pull and connectivity determinants for intra-European migration was analysed in three steps. Pearson correlation coefficients were used to establish the important explanatory variables. Factor analysis was used to deal with potential multi-co-linearity and reduce the large number of explanatory variables to a smaller number of factors. Finally, factor scores were used as inputs to linear regression models.

### 3.2. Correlation Analysis

It was found that 8 and 13 of the 22 Pearson correlation coefficients existing between the independent variables and the dependent variables 1 and 2 were higher than 0.1 and significant at the 0.05 level in periods 1997 – 2004 and 2005 – 2013 respectively (Table A1):

- Dependent variable 1 follows the geographical dispersal of migration stocks only. The variables representing economic connectivity (trade), knowledge

connectivity (patents), tourism connectivity (nights spent) and language connectivity generated Pearson coefficients above 0.1 for both periods. Both traditional push-pull variables (income and employment disparities) and the non-monetary costs and returns of migration turned out to be insignificant for both periods. Most migrant exchange happens within the 'rich club' members of Europe. The European centre countries account for relatively homogenous economic and social environments and generate relatively low potential for income-driven migration. Moreover, flows by labour migrants from south to north are often counterbalanced by flows of life-style migrants from north to south.

- Dependent variable 2 represents both the structure and the intensity of migration flows. For the period 1997 – 2004, the connectivity variables 14 – 21 (trade, patents, nights spent and language skills) proved significant and generated Pearson coefficients above 0.1. The push-pull variables generated quite low correlation coefficients for this period, 1997 – 2004. The disparities in wages, employment level and quality of life determinants were too low to generate intense migration flows among the EU-15 member countries. The situation changed after the Eastern enlargement in 2004. Income disparities between the East and West were substantial and generated vast migration flows. The non--monetary costs and benefits of migration also seemed important for this intensity of migration flow. All correlation coefficients have the signs which were expected for them in Table A1. Correlations coefficients for the wage disparities, for the ESS variables for self-reported levels of personal happiness, for satisfaction with the current economic performance of the country, for life satisfaction and opinions on the state of the country's democracy are all negative, indicating that Europeans prefer migrating to countries with higher wages and a better quality of life.

The correlation analysis indicated that connectivity factors show stronger associations with intra-European migration flows than do traditional push and pull factors when the intensity of flows is taken into account. When the 'pure' structure of destinations is taken into account, the connectivities in tourism, trade and knowledge flows are significantly correlated with the geographical structure of the intra-European migrant stocks. Also important are language connectivities. These reflect both globalities in the form of centres of gravity (Mahroum 2000) in respect of economic variables, and regional spatialities in terms of language. As for the traditional push-pull factors, wage disparities turned out to be important for explaining changes in migrant stocks after the EU's Eastern enlargement. Disparities in GDP levels, unemployment levels and welfare payments, on the other hand, proved insignificant in terms of motivating the intra-European migration flows in both periods.

### 3.3. Factor Analysis and Regressions

We have used a high number of independent variables in order to explain variations in the structures and/or intensities of migration flows. Some independent variables are mutually inter-correlated. Merchandise imports and exports, for example, generate Pearson coefficients of 0.854 in 1997 – 2004 and 0.872 in 2005 – 2013. The factor analysis was used to deal with the multi-co-linearity amongst the independent variables, and in order to reduce these to a smaller number of factors:

- For the period 1997 – 2004, the connectivity variables regarding merchandise trade (nos. 14 and 15), patents (nos. 16 – 18) and nights spent (no. 19) loaded on Factor 1 ('Connectivity'). Two language variables (nos. 20 and 21) loaded on Factor 2 ('Languages'). The language variables are correlated, because the language which is considered useful usually overlaps with the first or second language known (except for the mother language). Factors 1 and 2 explained 47.6% and 23.9% of the total variance for dependent variables 1 and 2 respectively in the period 1997 – 2004. The composition of factors was identical for both dependent variables (Table A2).
- As for the period 2005 – 2013, the same connectivity and language variables generated almost identical Factors 1 and 2 ('Connectivity' and 'Languages') for dependent variable 1. Factors 1 and 2 explained 47.1% and 27.9% of the total variance respectively in period 2005 – 2013 (Table A3). Moreover, two variables on wage disparities (nos. 2 and 3), and four ESS variables (nos. 8, 9, 10 and 22) combined in Factor 3 ('push-pull'). Factors 1, 2 and 3 explained 27.8%, 13.8% and 36.7% of the total variance respectively in variable 2 in the period 2005 – 2013 (Table A4).

Both the Kaiser-Olkin-Mayer (KMO) test and the Bartlet sphericity tests confirmed that the factor analysis produced satisfactory solutions (see Tables A.2, A.3 and A.4 in the Appendix).

The factor scores were used as inputs to the linear regressions which were run for both of the time periods: 1997 – 2004 and 2005 – 2013 (Table 3):

- The regressions explained 49.1% and 31.9% of the total variance in the dependent variable 1 in the periods 1997 – 2004 and 2005 – 2013 respectively. These levels are relatively high given the significant diversity of the intra-European migration flows. The diversity of migration motives increased in the period 2005 – 2013 as compared with that of the previous period – after the new member countries obtained access to the labour and knowledge markets of Western Europe. Labour migrants, welfare seekers, but also tertiary students and life-style migrants were behind substantial increases in total migrant stocks in Europe. Increased diversity of migration motives was likely to be behind decreases in the

total explanatory power of regression in the period 2005 – 2013 as compared to 1997 – 2004.

• The regressions explained 18.0% and 16.6% of the total variance in the dependent variable 2 in the periods 1997 – 2004 and 2005 – 2013 respectively. Dependent variable 2 is more complex than dependent variable 1, as it represents both the intensity and the geographical structure of migration flows. Variable 2 therefore possesses generally lower levels of explanatory power than variable 1.

T a b l e 3

**Linear Regression of Push-pull and Connectivity Factors**

| | **1997 – 2004** | | | | | | **2005 – 2013** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Variable 1** | | | **Variable 2** | | | **Variable 1** | | | **Variable 2** | | |
| | **Beta** | ***t*** | **Sig** | **Beta** | ***t*** | **Sig** | **Beta** | ***t*** | **Sig** | **Beta** | ***t*** | **Sig** |
| Constant | | 15.338 | 0.000 | | 5.474 | 0.000 | | 14.456 | 0.000 | | 9.087 | 0.000 |
| F1 Connectivity | 0.669 | 10.651 | 0.000 | 0.379 | 4.759 | 0.000 | 0.476 | 6.748 | 0.000 | 0.297 | 3.788 | 0.000 |
| F2 Languages | 0.207 | 3.292 | 0.001 | 0.189 | 2.372 | 0.019 | 0.305 | 4.334 | 0.000 | 0.150 | 1.912 | 0.058 |
| F3 Push-pull | -- | -- | -- | -- | -- | -- | -- | -- | -- | 0.235 | –3.005 | 0.003 |
| Adjusted $R^2$ | **0.491*** ** | | | **0.180*** ** | | | **0.319*** ** | | | **0.166*** ** | | |

*Notes*: B – standardised regression coefficient. *** significant on 0.001 level.

*Source*: Authors' computations.

The standardized regression coefficient (Beta) indicates which of the independent variables have a greater effect on the dependent variable in a multiple regression analysis – when the variables are measured in different units of measurement:

• Factor 1, connectivity, consistently had the highest Beta values in both time periods. It remained the strongest predictor of intra-European migrant stocks (Hypothesis 2). The relative importance of Factor 1 decreased over time in relation to both dependent variables 1 and 2 (Table 3). This decrease is probably related to both the territorial re-orientation of intra-Europeans flows in tourism, trade and knowledge (independent variables), and to the re-orientation of migrant flows (dependent variables). The decrease in the relative importance of Factor 1 also refers to the increased diversity in the intra-European migration system after 2004.

• Factor 2, language, decreased in importance between the two time periods in relation to dependent variable 2 (structure and intensity of migration flows). The same factor, however, increased in importance in relation to dependent variable 1 (structure of migration flows). This indicates that while many high-intensity migration flows developed between countries speaking different languages (e.g. Romania to Italy and Spain), there also was an increase in flows related

to language similarity (e.g. Germany to Switzerland and Austria). Factor 2 also embodies the growing importance of English as a global language ('language known' and 'language useful', Hypothesis 3).

- Factor 3, push-pull variables, proved significant only for the period 2005 – 2013 as related to dependent variable 2. This is to be expected, as this factor essentially captures the vast increase in the stock of migrants from the new member countries moving to Western Europe (Hypothesis 1).

## 4. Discussion and Conclusions

Both the network analysis and the factor and regression analysis support the idea of an intra-European migration system (Hypothesis 2). Such a system is an identifiable geographical structure that persists across space and time. A migration system is a product of interacting nation-states and corresponding socio-cultural, geopolitical, and economic factors and policies (Zlotnik, 1999; DeWaard, Kim and Raymer, 2012). The world migration system accounts for the high diversity which exists in terms of origins and destinations, yet the relatively small number of countries which account for three quarters of immigrant intake. We found the same pattern for Europe: six countries accounted for the intake of over 3/4 of all intra-European migrants in 1997 – 2013. Zlotnik's findings on the concept of a world-wide, relatively stable system of international migrations were corroborated in our study also for the system of intra-European migration.

Stability of the network does not imply that the network of intra-European migration is static; on the contrary, the network includes dynamic relationships between countries of origin and destination. Neighbourhood and (broad) language-proximity, for example, significantly informed memberships of individual modules. The dynamic nature of the intra-European migration networks is demonstrated by the emergence of new sets of institutions shaping migration flows (visa-free travel, opening labour markets, student mobility programmes, and the introduction of the new transport modes). The UK-centred migrant inflows from Poland and other Eastern EU members, for example, are not informed by the traditional neighbour/language proximity framework, but by the rising importance of English as global language, and by the availability of low-cost travel (Jenissen, 2007).

Our study has some important limitations. The first, and most obvious one, is related to the availability and quality of the data on international migrant stocks and flows. Migration research lacks comprehensive and harmonised data on migrant stocks in Europe. There are different concepts of migrant stocks. Some countries report stocks of foreign nationals by country of birth, which therefore

includes nationals who have already acquired destination country citizenship. Other countries apply the concept of 'stock of foreign nationals by nationality' instead. While considerable research exists on most migrant groups (King, 2012), this is fragmented and selective (Castles, 2010) – focusing on migrants versus refugees/asylum seekers, short versus long term versus 'permanent' migration, and specific demographic, social and national groups. Most data on migrant stocks comes from the labour force surveys and population registers. There are EU-citizens, mainly from Central and Eastern Europe, who stay in EU-15 countries but fail to register with the national population or tax registers. Good quality data on migrant stocks are missing for some European countries (Ukraine, Serbia, Bosnia and Albania). Further research on long-term migration may do well to consider new data sources, such as (a) national social and health insurance institutions' data on foreign nationals; (b) data on car registrations by foreign nationals; and (c) data on bank accounts in host countries.

The second limitation is related to the changing nature of contemporary international migration. The removal of visa barriers and the rise of low-cost travel increasingly blurs the borders between long-term and short-term, and permanent and circular migration. As for the development of the migration network, the EU Eastern enlargements (in 2004 and 2007) were a slow and uneven process. Restrictions in travel and restrictions in access to the EU-15 labour market were removed gradually over 2004 – 2014. We assume here that the removal of travel restrictions was more important in relation to changes in migrant stocks than the removal of restrictions on official access to labour markets. Many EU-10 migrants lived and worked in EU-15 member countries prior to the official opening of labour markets. The blurred boundary between the official and the actual opening of labour markets represents an important limitation on our research.

Contemporary intra-European migration and mobility flows are accounted for by the high diversity of motives and duration patterns. There is a substantial lack of data on specific migrations and mobilities (e.g. retirement migration, life-style migration, circular migration). Return and circular migration are especially important in post-enlargement Europe. The Eastern European migrants, for example, are more likely to engage in temporary circular (serial) and transnational mobility (Favell, 2008; Martin and Radu, 2012). Migration researchers have responded by devising new concepts which seek to describe the more diverse reality of migration pathways – incomplete migration (Bonifazi et al., 2014), free movers (Favell, 2008), liquid migration (Engbersen, Snel and De Boom, 2010), circular and return migration flows, and multiple migrations (Ciobanu, 2015). The difficulty has been in operationalizing these concepts, particularly in relation to quantitative data.

For mapping mobility patterns; (i) data provided by traditional tourism/passenger surveys; and (ii) telecom operator data on roaming services could be useful resources, assuming that short-term visitors use phone numbers registered in their country of origin.

## References


ALVAREZ-PLATA, P. – BRÜCKER, H. – SILIVERSTOVS, B. (2003): Potential Migration from Central and Eastern Europe into the EU-15 – An Update. [Report for the European Commission, DG Employment and Social Affairs.] Berlin: Deutsches Institut für Wirtschaftsforschung.

ANDERSON, J. (2015): Labour Migration and International Trade. Journal of Economics and Econometrics, *58*, No. 3, pp. 1 – 29.

BAKEWELL, O. (2014): Relaunching Migration Systems. Migration Studies, *2*, No. 3, pp. 300 –318.

BARRAT, A. – BARHTELEMY, M. – PASTOR-SATTORAS, R. – VESPIGNANI, A. (2004): The Architecture of Complex Weighted Networks. Proceedings of the National Academy of Sciences, *101*, No. 11, pp. 3747 – 3752.

BONIFAZI, C. – OKÓLSKI, M. – SCHOORL, J. – SIMON, P. (2014): International Migration in Europe. Amsterdam: Amsterdam University Press.

BORJAS, G. J. (1987): Self-selection and the Earnings of Immigrants. American Economic Review, *77*, No. 4, pp. 531 – 553.

BORJAS, G. J. (1999): Immigration and Welfare Magnets. Journal of Labor Economics, *17*, No. 4, pp. 607 – 637.

CASTLES S. (2010): Understanding Global Migration: A Social Transformation Perspective. Journal of Ethnic and Migration Studies, *36*, No. 10, pp. 1565 – 1586.

CASTRO-MARTÍN, T. – CORTINA, C. (2015): Demographic Issues of Intra-European Migration: Destinations, Family and Settlement. European Journal of Population, *31*, No. 2, pp. 109 – 125.

CIOBANU, R. O. (2015): Multiple Migration Flows of Romanians. Mobilities, *10*, No. 3, pp. 466 – 485.

De JONG, G. F. – CHAMRATRITHIRONG, A. – TRAN, Q. G. (2002): For Better, for Worse: Life Satisfaction Consequences of Migration. International Migration Review, *36*, No. 3, pp. 838 – 863.

DENNETT, A. – WILSON, A. (2013): A Multilevel Spatial Interaction Modelling Framework for Estimating Interregional Migration in Europe. Environment and Planning A, *45*, No. 6, pp. 1491 – 1507.

DeWAARD, J. – KIM, K. – RAYMER, J. (2012): Migration Systems in Europe: Evidence from Harmonized Flow Data. Demography, *49*, No. 4, pp. 1307 – 1333.

EASLEY, D. – KLEINBERG, J. (2010): Networks, Crowds, and Markets Reasoning About a Highly Connected World. Cambridge: Cambridge University Press.

ENGBERSEN, G. – SNEL, E. – De BOOM, J. (2010): A Van Full of Poles. Liquid Migration in Eastern and Central European Countries. In: BLACK, R. – ENGBERSEN, G. – OKÓLSKI, M. – PANŢÎRU, C. (eds): A Continent Moving West? EU Enlargement and Labor Migration from Central and Eastern Europe. Amsterdam: Amsterdam University Press, pp. 115 – 140.

ESS (European Social Survey) (2002 – 2012): Data and Documentation by Round/Year, Rounds 1 – 6. Available at: <http://www.europeansocialsurvey.org/data/round-index.html>.

EUROPEAN COMMISSION (2006): Europeans and their Languages – Special Eurobarometer 243. European Commission, February 2006.

EUROPEAN COMMISSION (2012): Europeans and their Languages – Special Eurobarometer 386. European Commission, June 2012.

EUROSTAT (2016): Database. Available at: <http://ec.europa.eu/eurostat/data/database>.

FAGIOLO, G. – MASTRORILLO, M. (2013): International Migration Network: Topology and Modelling. Physical Review E, *88*, No. 1, pp. 1 – 33.

FAVELL, A. (2008): The New Face of East-West Migration in Europe. Journal of Ethnic and Migration Studies, *34*, No. 5, pp. 701 – 716.

FAWCETT, J. T. (1989): Networks, Linkages, and Migration Systems. International Migration Review, *23*, No. 3, pp. 671 – 680.

FRUCHTERMAN, T. M. J. – REINGOLD, E. M. (1991): Graph Drawing by Force-Directed Placement. Software: Practice and Experience, *21*, No. 11, pp. 1129 – 1164.

GALGÓCZY, B. – LESCHKE, J. – WATT, M. A. (2009): Intra-EU Labour Migration – Flows and Policy Responses. In: GALGÓCZY, B. – LESCHKE, J. – WATT, M. A. (eds): EU Labour Migration since Enlargement: Trends, Impacts and Policies. Aldershot, UK: Ashgate, pp. 1 – 28.

GENC, M. – GHEASI, M. – NIJKAMP, P. – POOT, J. (2011): The Impact of Immigration on International Trade: A Meta-Analysis. [Discussion Paper, No. 6145.] Bonn: Institute for the Study of Labor (IZA).

GIULIETTI, C. – GUZI, M. – KAHANEC, M. – ZIMMERMANN, K. F. (2013): Unemployment Benefits and Immigration: Evidence from the EU. International Journal of Manpower, *34*, No. 1, pp. 24 – 38.

HARRIS, J. – TODARO, M. P. (1970): Migration, Unemployment and Development: A Two-Sector Analysis. American Economic Review, *60*, No. 1, pp. 126 – 142.

HOOGHE, M. – TRAPPERS, A. – MEULEMAN, B. – REESKENS, T. (2008): Migration to European Countries: A Structural Explanation of Patterns, 1980 – 2004. International Migration Review, *42*, No. 2, pp. 476 – 504.

JENISSEN, R. (2007): Causality Chains in the International Migration Systems Approach. Population Research and Policy Review, *26*, No. 4, pp. 411 – 436.

KAHANEC, M. – PYTLIKOVÁ, M. – ZIMMERMANN, K. F. (2016): The Free Movement of Workers in an Enlarged European Union: Institutional Underpinnings of Economic Adjustment. In: KAHANEC, M. – ZIMMERMANN, K. F. (eds): Labor Migration, EU Enlargement and the Great Recession. Berlin: Springer, pp 1 – 34.

KAHANEC, M. – ZIMMERMANN, K. F. (2016): EU Post-enlargement Migration and the Great Recession: Lessons and Policy Implications, In: KAHANEC, M. – ZIMMERMANN, K. F. (eds): Labor Migration, EU Enlargement and the Great Recession. Berlin: Springer, pp. 419 – 445.

KING, R. (2002): Towards a New Map of European Migration. International Journal of Population Geography, *8*, No. 2, pp. 89 – 106.

KING, R. (2012): Geography and Migration Studies: Retrospect and Prospect. Population, Space and Place, *18*, No. 2, pp. 134 – 153.

LANDESMANN, M. – LEITNER, S. M. (2015): Intra-EU Mobility and Push and Pull Factors in EU Labour Markets: Estimating a Panel VAR Model. [Working Paper 120.] Vienna: Vienna Institute for International Economic Studies.

MAHROUM, S. (2000): Highly Skilled Globetrotters: Mapping the International Migration of Human Capital. R&D Management, *30*, No. 1, pp. 23 – 32.

MARTIN, R. – RADU, D. (2012): Return Migration: The Experience of Eastern Europe. International Migration, *50*, No. 6, pp. 109 – 128.

MAYDA, A. M. (2010): International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. Journal of Population Economics, *23*, No. 4, pp. 1249 – 1274.

MINCER, J. (1978): Family Migration Decision. Journal of Political Economy, *86*, No. 5, pp. 749 – 773.

NEWMAN, M. E. J. (2006): Modularity and Community Structure in Networks. Proceedings of the National Academy of Sciences of the United States of America, *103*, No. 23, pp. 8577 – 8696.

OECD (2016): OECD. Stat. Available at: <http://stats.oecd.org/>.

OLGIATI, A. – CALVO, R. – BERKMAN, L. (2013). Are Migrants Going Up a Blind Alley? Economic Migration and Life Satisfaction around the World: Cross-national Evidence from Europe, North America and Australia. Social Indicators Research, *114*, No. 2, pp. 384 – 404.

PARUTIS, V. (2014): "Economic Migrants" or "Middling Transnationals"? East European Migrants' Experiences of Work in the UK. International Migration, *52*, No. 1, pp. 36 – 55.
PEDERSEN, P. J. – PYTLIKOVÁ, M. – SMITH, N. (2008): Selection or Network Effects? Migration Flows into 27 OECD Countries. European Economic Review, *52*, No. 7, pp. 1160 – 1186.
SARDADVAR, S. – ROCHA-AKIS, S. (2016): Interregional Migration within the European Union in the Aftermath of the Eastern Enlargements: A Spatial Approach. Review of Regional Research, *36*, No. 1, pp. 51 – 79.
SJAASTAD, L. A. (1962): The Costs and Returns of Human Migration. The Journal of Political Economy, *70*, No. 5, pp. 80 – 93.
VERWIEBE, R. – WIESBÖCK, L. – TEITZER, R. (2014): New Forms of Intra-European Migration, Labour Market Dynamics and Social Inequality in Europe. Migration Letters, *11*, No. 2, pp. 125 – 136.
ZLOTNIK, H. (1999): Trends of International Migration since 1965: What Existing Data Reveal. International Migration, *37*, No. 1, pp. 21 – 61.

# A p p e n d i c e s

## Appendix 1: Data Sources and Limitations

As for the dependent and independent variables, the annual averages for the periods 1997 – 2004 and 2005 – 2013 were used. Annual averages were used in order to cope with missing data for particular years for some independent variables

Stocks of foreign populations (dependent variables) are provided by the OECD International Migration Database. The basic variable is 'stock of foreign nationals by country of birth'. Where country of birth data is missing (e.g. for Germany, Sweden, and Italy), 'stocks of foreign nationals by nationality' provides a surrogate measure. France and Austria provide only one or two data points within the 10 years, and these were treated as period averages. In the case of some less important sending countries, there are no data at all on stocks, but only for flows, so it was necessary to obtain estimates of the former. A one-year timing criterion is used for reporting long-term migration.

The Eurostat (2016) provided data on GDP in purchasing power parity levels (variable 1), average wages (variables 2 and 3), unemployment rates (variables 4, 5 and 6) and social benefits (variable 7).

The European Social Survey (ESS) provided data for questions concerning satisfaction with various domains of private and public life: 'All things considered, how satisfied are you with your life as a whole nowadays?' (variable 8). 'On the whole how satisfied are you with the present state of the economy?' (variable 9). 'On the whole, how satisfied are you with the way democracy works?' (variable 10). 'What do you think overall about the state of education nowadays?' (variable 11). Answers for variables 17 – 20 ranged from 0 – 'extremely

dissatisfied' to 10 – 'extremely satisfied' (10). Two questions concerned personal opinions and attitudes on happiness and trust (variables 12 and 13): 'Taking all things together, how happy would you say you are?' Answers for variable 12 ranged from 0 – 'extremely unhappy' to 10 – 'extremely unhappy'. 'Generally speaking, would you say that most people can be trusted, or that you can't be too careful?' Answers for variable 13 ranged from 0 – 'you can't be too careful' to 10 – 'most people can be trusted'. National averages were computed for variables 8 – 13. Differences in averages were used as inputs to regression. Data from Round 1 (ESS, 2002) were used for the period 1998 – 2002, data from Round 3 (2006) for the period 2003 – 2007 and data from Round 6 (2012) for the period 2008 – 2012. All the above-mentioned questions were present in all ESS rounds.

The Eurostat database (2016) provided data on the structure of merchandise imports and exports by country of origin (variables 14 and 15). The OECD database (2016) provided data for the three variables on patents (16, 17 and 18) based on the general relationship between inflows/outflows and stocks of foreign nationals. Data on tourism (nights spent, variable 19) are provided by Eurostat for 1998 – 2013, and where data are missing for particular years, these were estimated as the average numbers of nights based on the data for previous and/or following years.

The special European Barometers 243 and 386 (European Commission, 2006; 2012) provided data on language skills (language known and language useful, variables 20 and 21). The question for variable 20 asked: 'Which languages do you speak well enough in order to be able to have a conversation, excluding your mother tongue?' The question for variable 21 asked 'Which two languages, apart from your mother tongue do you think are the most useful to know for your personal development and career?' Averages of the ESS rounds 1 and 2 for period 1997 – 2004 and averages of the rounds 3, 4, 5 and 6 for period 2005 – 2013 were used for computation of the psychological costs.

We used the percentages of languages identified by home country population for the top 5 migrant destinations. Driving distances for European capitals (variable 22) are according to the <http://www.travelnotes.org/Europe/Distances/index.htm>.

## Appendix 2: Correlation Analysis

T a b l e A1

**Pearson Correlations between the Determinants and Student Outflows, by Dependent Variable and Time Period**

| | Variable 1 | | | | Variable 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | 1997 – 2004 | | 2005 – 2013 | | 1997 – 2004 | | 2005 – 2013 | |
| | Pearson | Sig. | Pearson | Sig. | Pearson | Sig. | Pearson | Sig. |
| ***Economic push-pull variables*** | | | | | | | | |
| 1. GDP (PPS) levels | –0.025 | 0.761 | 0.004 | 0.961 | 0.042 | 0.606 | –0.146 | 0.070 |
| 2. Average wage (single. no children) | –0.020 | 0.823 | –0.031 | 0.705 | 0.031 | 0.721 | **–0.213** | **0.008** |
| 3. Average wage (married. two children) | –0.028 | 0.744 | –0.037 | 0.651 | 0.027 | 0.761 | **–0.217** | **0.007** |
| 4. Social benefits | –0.106 | 0.206 | 0.054 | 0.501 | 0.012 | 0.889 | –0.090 | 0.264 |
| 5. Unemployment rate total | –0.120 | 0.138 | –0.027 | 0.737 | –0.079 | 0.327 | 0.047 | 0.557 |
| 6. Unemployment rate (up to age 25) | 0.002 | 0.976 | 0.031 | 0.701 | –0.008 | 0.923 | 0.091 | 0.261 |
| 7. Long-term unemployment rate | –0.134 | 0.097 | –0.047 | 0.563 | –0.074 | 0.360 | 0.041 | 0.615 |
| ***Non-monetary costs and benefits (ESS)*** | | | | | | | | |
| 8. Life satisfaction | 0.033 | 0.687 | –0.038 | 0.637 | 0.064 | 0.425 | **–0.172** | **0.032** |
| 9. Satisfaction with current econ. Performance | 0.147 | 0.067 | –0.128 | 0.113 | 0.045 | 0.580 | **–0.161** | **0.045** |
| 10. Opinions on the state of democracy | 0.036 | 0.659 | –0.117 | 0.148 | 0.032 | 0.690 | **–0.192** | **0.017** |
| 11. Satisfaction with quality of education | 0.106 | 0.188 | –0.018 | 0.825 | 0.059 | 0.462 | –0.142 | 0.079 |
| 12. Self-reported levels of personal happiness | –0.004 | 0.960 | –0.015 | 0.852 | 0.064 | 0.425 | **–0.200** | **0.013** |
| 13. Self-reported levels of personal trust | 0.005 | 0.949 | 0.007 | 0.926 | 0.022 | 0.788 | –0.148 | 0.066 |
| ***Connectivity variables*** | | | | | | | | |
| 14. Merchandise imports shares | **0.643** | **0.000** | **0.456** | **0.000** | **0.457** | **0.000** | **0.260** | **0.002** |
| 15. Merchandise exports shares | **0.625** | **0.000** | **0.456** | **0.000** | **0.333** | **0.000** | **0.254** | **0.002** |
| 16. Foreign ownership of domestic patents | **0.475** | **0.000** | **0.258** | **0.001** | **0.171** | **0.033** | 0.134 | 0.096 |
| 17. Domestic ownership of foreign patents | **0.381** | **0.000** | **0.323** | **0.000** | **0.281** | **0.000** | **0.158** | **0.050** |
| 18. Patents with foreign co-inventor(s) | **0.476** | **0.000** | **0.366** | **0.000** | **0.240** | **0.003** | **0.291** | **0.000** |
| 19. Nights spent by foreign tourists | **0.664** | **0.000** | **0.540** | **0.000** | **0.387** | **0.000** | **0.289** | **0.000** |
| 20. Language known | **0.291** | **0.000** | **0.338** | **0.000** | **0.233** | **0.003** | **0.154** | **0.050** |
| 21. Language useful | **0.217** | **0.007** | **0.295** | **0.000** | **0.173** | **0.032** | **0.178** | **0.026** |
| 22. Driving distance between capitals | –0.155 | 0.054 | –0.105 | 0.192 | –0.025 | 0.761 | 0.070 | 0.387 |

*Notes*: Correlations over 0.1 and significant at the 0.05 level in bold and shading.

*Source*: Authors' computations.

## Appendix 3: Factor Analysis

T a b l e A2

**Factor Analysis for Variable 1 and Variable 2 in Period 1997 – 2004**

| | **Component** (total variance explained) | |
|---|---|---|
| | **1** (47.6%) connectivity | **2** (23.9%) languages |
| Merchandise import shares | **0.868** | 0.088 |
| Merchandise export shares | **0.859** | 0.070 |
| Patents – co-authorship | **0.819** | 0.101 |
| Patents foreign ownership | **0.790** | 0.009 |
| Nights spent | **0.767** | 0.063 |
| Patents domestic ownership | **0.642** | 0.343 |
| Language known | 0.136 | **0.948** |
| Language useful | 0.042 | **0.933** |

*Note:* Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization Rotation converged in 3 iterations. Kaiser-Meyer-Olkin test = 0.722. Bartlet's test of sphericity: Approx. Chi-Square = 706.866; df = 28; Sig = 0.000. Total variance explained: 71.5%.

*Source*: Authors' computations.

T a b l e A3

**Factor Analysis for Period 2005 – 2013, Model for Variable 1**

| | **Component** (total variance explained) | |
|---|---|---|
| | **1** (47.1%) connectivity | **2** (27.9%) languages |
| Merchandise import shares | **0.882** | 0.029 |
| Merchandise export shares | **0.874** | –0.011 |
| Patents – co-authorship | **0.840** | 0.107 |
| Patents domestic ownership | **0.750** | 0.255 |
| Nights spent | **0.736** | 0.216 |
| Patents foreign ownership | **0.618** | 0.047 |
| Language useful | 0.036 | **0.915** |
| Language known | 0.175 | **0.888** |

*Note*: Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization Rotation converged in 3 iterations. Kaiser-Meyer-Olkin test = 0.769. Bartlet's test of sphericity: Approx. Chi-Square = 620.040; df = 78; Sig = 0.000. Total variance explained: 69.1%.

*Source:* Authors' computations.

T a b l e A4

**Factor Analysis for Period 2005 – 2013, Model for Variable 2**

| | **Component** (total variance explained) | | |
|---|---|---|---|
| | **1** (27.8%) connectivity | **2** (13.8%) languages | **3** (36.7%) push-pull |
| Average wage married 2 children | 0.102 | –0.036 | **0.921** |
| Average wage single no children | 0.123 | –0.036 | **0.914** |
| ESS Satisfaction with democracy | –0.107 | –0.069 | **0.912** |
| ESS Life satisfaction | 0.046 | 0.194 | **0.908** |
| ESS Happiness | 0.107 | 0.272 | **0.856** |
| ESS Satisfaction with economic performance | –0.277 | –0.005 | **0.787** |
| Merchandise import shares | **0.905** | –0.012 | 0.048 |
| Merchandise export shares | **0.897** | –0.044 | 0.067 |
| Patents – Co-authorship | **0.808** | 0.091 | –0.150 |
| Patents domestic ownership | **0.777** | 0.200 | 0.170 |
| Nights spent | **0.742** | 0.207 | –0.101 |
| Language useful | 0.053 | **0.920** | –0.003 |
| Language known | 0.221 | **0.858** | 0.142 |

*Note:* Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization Rotation converged in 4 iterations. Kaiser-Meyer-Olkin test = 0.741. Bartlet's test of sphericity: Approx. Chi-Square = 2307.254; df = 78; Sig = 0.000. Total variance explained: 78.3%.

*Source*: Authors' computations.