Co-financed by the European Regional Development Fund

Inspire Policy Making with Territorial Evidence

FINAL REPORT //

Interregional migration flows in Europe

IRiE – Interregional Relations in Europe

Annex 5 // October 2021

This FINAL REPORT is conducted within the framework of the ESPON 2020 Cooperation Programme, partly financed by the European Regional Development Fund.

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway, Switzerland and the United Kingdom.

Coordination:
Nicolas Rossignol, Head of Unit for Evidence and Outreach (ESPON EGTC); Marta Roca, Financial Expert (ESPON EGTC); Xabier Velasco Echeverría, Head of Territorial Observatory of Navarra (NASUVINSA, lead partner)

Communication:
Nikos Lampropoulos, Project Expert - Press and Media Activities (ESPON EGTC); Sheila Izquieta Rojano, NASUVINSA (Spain)

Authors
Tomasz Komornicki – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland); Wojciech Pomianowski – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland); Rafał Wiśniewski – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland); Barbara Szejgiec-Kolenda – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland); Marcin Mazur – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland); Dorota Ciołek, University of Gdansk, Faculty of Management; Patryk Duma – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland).

Technical Support
Marta Zabost – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland); Julia Wójcik – Institute of Geography and Spatial Organization, Polish Academy of Sciences (Poland).

Advisory group
Marie Lorraine Dangeard, National Agency for Territorial Cohesion, France; Margarita Golovko, Regional Development Department, Ministry of Finance, Estonia; Ivan Lajtman, Directorate for Regional Development, Ministry of Regional Development and EU Funds, Croatia.

Acknowledgements
Andrea Conte (REMO Team, JRC Seville); Giovanni Mandras (REMO Team, JRC Seville)

Information on ESPON and its projects can be found at www.espon.eu.
The website provides the possibility to download and examine the most recent documents produced by finalised and ongoing ESPON projects.

Layout and graphic design by BGRAPHIC, Denmark

Contact: info@espon.eu

Co-financed by the European Regional Development Fund

Inspire Policy Making with Territorial Evidence

FINAL REPORT //

Interregional migration flows in Europe

IRiE – Interregional Relations in Europe

Annex 5 // October 2021

Table of contents

List of figures and tables

List of figures

List of tables

Abbreviations

C2C	Country to country
EU	European Union
EUROSTAT	European Statistical Office
GDP	Gross Domestic Product
IRiE	Interregional Relations in Europe
NSI	National Statistic Institution
NUTS2	Nomenclature of Territorial Units for Statistics
Pop	Population-based outflow
R2R	Region to region
TSE	Time series extension

Abstract

The present report contains the results of our research into migration flows at two spatial levels, national and regional (NUTS 2), whose effect is constituted by the complete matrices of flows. Both external (international) and internal (domestic) migrations are accounted for.

The primary objective of the study was to determine the magnitudes of the interregional flows and their directions within the ESPON space countries. Other objectives included the identification of factors that influence international migration flows at the regional level, recognition of the new spatial structures in the regional perspective, and the respective typology of the regions.

The methodology chapter describes in detail the entire data-acquisition procedure and the methodological procedures used in case of data gaps in the C2C and R2R matrices. The biggest challenge at this stage was to create a complete matrix of regional flows, especially in terms of intra-country flows.

Chapter 3 presents the most important results of the work on migration flows, distinguishing migrations at the country and regional levels. The dominant direction of connections in Europe in the years 2010-2018 was East-West (visible at both the spatial levels analysed). These flows encompassed migration movements (including return and pendulum movements) initiated after the accession in 2004 to the EU of ten new member countries. In addition, migrations also resulted from the global financial crisis, which began in 2008. This concerned an intensification of the migration outflow from Spain, Portugal, and Greece — countries that in the years 2010-2018 exhibited a negative total migration balance. Flows at the regional level are presented in five dimensions: intensity (size), connectivity, balance (comparison of outflow and inflow), concentration (dispersion of senders and receivers), and distance (how far people migrate — to what extent distance determines the size of flows). Synthesis of the indicators has been presented (typologies): the indicator for migration-flow intensity has been set against the indicators for balance, concentration, and migration-flow distance. This chapter ends with an analysis of explanatory factors.

Keywords

migrant, migration flow, regional migration

Highlights

Data

- Available data on migration flows are inconsistent between countries of origin and destination even when migrants are counted by the common EUROSTAT definition.
- Two comprehensive sources for data on migration flows within the ESPON space at national and regional level (EUROSTAT and NSI) have been used, but both are incomplete.
- The gaps in migration flows have been filled on the basis of estimation measures based on migrant stock, population, and GDP.

Methodology

- To obtain complex information on spatial and temporal allocation of migration flows within ESPON Space at the national level (C2C) we have used a procedure including the following steps: Base Data, Stock Gain estimation, In-Out-Cross estimation.
- The raw data was organised into a 32×32 C2C matrix format, with countries of origin in rows and countries of destination in columns. This involved transposition for data reported at destination. The result was two alternative sets of 10 one-year matrices of 992 cells each. In each case, the raw data could provide no more than two-thirds of the total number of cells required.
- Another set of C2C cells was estimated on the basis of the stock of migrants (permanent residents of foreign origin). It was assumed that new migrants followed an established spatial migration pattern.
- After application of the Stock-Gain method, missing cells still existed for the following countries: CZ, IE, EL, ES, CV, LV, MT, PL, PT, RO, UK. For these countries, we established two alternative linear models (In-Out-Cross estimation):

  - An Outflow model, based on population as an explanatory variable

  - An Inflow model, based on GDP as explanatory variable
- The estimation of regional level migration (R2R) was largely determined by the existence of C2C flows. Our task was to decompose the known number of migrants into spatial units, rather than estimate the number of migrants from scratch, using the population dynamics and structure of each spatial unit separately.
- The R2R harmonization guideline can be summarized as follows: Migrants adhere to a country-wide pattern at the source and imitate previous migrants at the destination. We believe this guideline is relatively safe and conservative, and thus best suited to incomplete and risky data sources.

The input to region-to-region flows estimation consisted of three data structures:

- C2C Migration Flows Matrix (generated as described in section C2C flows)

- R-Outflow table of Regional Migration Outflow as Outflow Driver

- R-Stock table of Regional Distribution of foreign population (Stock) as Inflow Driver

New territorial evidence

- The dominant direction of migration-related connections in the ESPON space was determined by East-West flows, i.e. those from the countries that joined the EU in 2004 (the biggest flows were from Poland and Romania).
- For 2010 Germany was clearly dominant for both inflows and outflows of migration. In the last year analysed (2018) the same countries still concentrated the majority of flows inside the ESPON space, but their respective shares had changed. Germany accounted for more than 25% of all the flows (with inflows exceeding outflows).
- The biggest positive net balance characterized Germany, the UK, and Switzerland, while the biggest negative balances were observed for Romania, Poland, and Spain.
- In 2018 the highest magnitude of flows (inflows and outflows jointly) in relation to the given region’s population occurred in the UK, where for a significant number of regions the value of this indicator exceeded 70 persons per 1,000 inhabitants.
- A decrease in the magnitude of migration outflows in the regional setting occurred first of all in the majority of the Spanish regions (e.g. Cataluña, Comunidad de Madrid), in the whole of Portugal, Ireland, and Lithuania (the highest value among all of the regions analysed), as well as in southwestern Poland (Śląskie, Dolnośląskie).
- There were characteristic differences in the shares of domestic migrations within some of the countries considered. They were visible in Poland (higher share of internal flows in the eastern part of the country), Spain (lower significance of such flows in the south – in the regions attracting foreign migrants), Germany (more international migration flows in western lands), and also Italy (domestic migrations more important in the south). It can be supposed that these differences resulted from such factors as: a) internal differences in development level and human capital; b) participation in the inter-metropolitan exchange of highly skilled personnel; c) high residential attractiveness.
- Most of the strongest migration flows in Europe were internal to the countries considered. The domination of domestic migrations may be interpreted as a persistence of the “border effect” in social relations, even between countries, where integration processes are strong and long lasting.

- Our study confirms the polycentric character of the German, British (primarily English), and Dutch settlement networks. Likewise, the connection networks in Sweden and Spain, as well as Romania, also turned out to be polycentric. On the other hand, migratory monocentricity was observed in France, Greece, and Bulgaria.
- From this perspective the spatial pattern of the migration core of Europe related partly to the so-called Blue Banana, although the intensity of flows distinctly decreased in the southern direction, with the values in northern Italy being perceptibly lower.
- Capital regions displayed a clear dominance of inflows over outflows. These migration surpluses occurred in the regions encompassing Warsaw (indicator value among the highest in Europe), Berlin, Budapest, Bucharest, Sofia, Vilnius, and Athens.
- Distinctly positive migration balance values were observed in the insular regions of Spain (Baleares, Canary Islands) and of France (Corsica). There was, however, no analogue in the Italian islands (Sicily, Sardinia), the Portuguese ones (Azores, Madeira), or the Greek ones.
- The migration balance of 2010-2018 very distinctly strengthened only a few of the concrete European MEGAs: Munich, Frankfurt, London, Berlin, and Warsaw.
- Our analysis seems to confirm an advancing concentration of demographic potential within the “Blue Banana”, Scandinavia, and a few separate areas in Central-Eastern Europe.
- The indicator reached its highest value — on the European scale — for Warsaw. This resulted from the multidirectional migration inflow from the territory of Poland, with simultaneous significant flow of outward foreign migrants. A similar situation, although on a smaller scale, was visible also in Bratislava, Prague, Budapest, and Bucharest. These cities can be considered migration “nodes” (links of the migratory movement chains).
- Domestic migrants, attracted to the metropolis, sometimes decided on further migration abroad. This was visible for Berlin, Warsaw, Prague, Bucharest, and Budapest, but also for Madrid, Helsinki, and Stockholm.
- Higher values of the indicator are observed in their western parts of the countries of Central-Eastern Europe (Czechia, Poland, Slovakia, Hungary). These may be associated with a concentration of migrations on shorter movements, either to domestic metropolises or to nearby Germany and Austria.
- In some cases (Central Europe) the dominant concentration on the inflow side can be associated with generally small and mostly neighbourhood immigration, with simultaneous, quite significant, and dispersed emigration (e.g. from eastern Germany to numerous NUTS 2 regions in the western part of the country).
- An essential observation is the diversification of the indicator values in regions characterized by a strong negative migration balance. Some of them feature a geographically

multidirectional outflow (in the regional perspective – e.g. southern Italy, western Poland, western Romania and Bulgaria, eastern Germany except for Brandenburg), without the domination of a single "receiver" region.

- Our overall assessment of the three typologies elaborated (in connection with our analysis of the basic indicators) indicates the separate character of some of the territorial clusters in terms of flow intensity, balance, concentration level, and distance of migrations.
- A few groups of geographically dispersed units (i.e. overseas territories, residentially attractive areas, metropolises of Central-Eastern Europe) had features in common.
- Our general model confirmed the significance of the associations between magnitude of migration flows and income levels as well as the wealth of the receiving regions (with the positive sign) and of the source regions (with the negative sign – gravity model).
- The affinity of languages (the very same language or the same group of languages) essentially enhanced the magnitude of migration flows.
- A bad labour market in the region of (potential) destination served as a barrier to migration flows.
- Regions of origin where a high percentage of the population was highly educated ought to have produced more intensive migration flows. The general model confirmed this hypothesis.
- Membership in the Schengen zone was a driver of migration flows, but the dependence was not so obvious for membership in the euro zone.
- Migrants' origin in countries having joined the EU after the year 2004 (in terms of the region of origin) was also a driver of migration flows.
- The opening up of the labour markets of particular regions stimulated migrations. The more recently a labour market had opened up, the greater its influence on migrations (novelty effect).
- The model for inflows to new member countries of the EU and non-metropolitan regions omits the significance of economic factors (GDP per capita and disposable income of households) in the destination region. Here an important spur to the flow of migrants was internal movements, frequently undertaken for non-economic reasons.
- Similarly, the economic criteria of GDP and household disposable income lost their significance in this model, which we established for outflows from regions of the "old" EU member countries.
- The model for internal (domestic) migrations did not show statistical significance for GDP per capita as a stimulant to migration flows (this is important for international migrations).

- For domestic migrations, the statistically significant factor turned out to be the migrants' belonging to metropolitan regions (in the regions of both origin and destination, the inter-metropolitan migrations of the staff, students, etc.).

1 Introduction

1.1 Background

Contemporary space is defined by a system of linkages and flows. The system corresponds to the historically determined distribution of socio-economic characteristics within territorial units. It depends on these characteristics and, at the same time, modifies them. In recent decades, the global processes that drive flows (relocation, geopolitical instability, economic inequalities, climate change) have been accelerating at an increasing rate. New forms of flows are being created (cf. e.g. Salt, 2008; Verwiebe et al., 2014), fostered by territorial and economic integration, as within the EU and the Schengen area (e.g. Davis, Gift, 2014). Without an analysis of international flows, it is impossible to describe spatio-functional structures accurately not only for countries, but also for regions or even local units. Many existing typologies of territorial units disregard their position in the space of flows. Thus, as globalisation and integration advance, the existing spatial delimitations lose their relevance. However, the system of linkages and flows is difficult to study (e.g. because of the inertia of official statistical systems). The resulting ignorance of the system leads to erroneous scientific diagnoses, and often also to misguided investments.

The need for wider, more comprehensive territorial research into linkages and flows was first noticed in theoretical studies (Castells 1998). Over time, opportunities and needs for empirical (policy-oriented) research were noticed at the European level (e.g. ESPON 1.4.3., ESPON FOCI) and within individual countries (in Poland e.g. Zaucha et al., 2014). The dynamics of particular types of flows — in particular their volatility associated with geopolitical and macroeconomic processes, such as economic transformation, the enlargement of the European Union, the 2008 economic crisis, the migration crisis, Brexit, global trade disputes, and the COVID-19 crisis — have proven to be particularly consequential.

Because of technological development and globalisation (including territorial integration) — which give rise, for example, to the emergence of new forms of mobility (inter alia, as a result of increased affordability of air transport) — flows of people, notably migratory, tourist, education-related flows, must be looked at from a different perspective. Never before have migration trends been as diverse as they are now (Castels and Miller, 2011), which indicates that migration processes are themselves globalising (see the concept of transnational social spaces: Pries, 1999; Faist, 2000). However, modern migrants “escape” the traditional definition of a migrant (in a way unrelated to the definition’s ambiguity). It is increasingly difficult to determine a primary place of residence (residing in several places is possible) or work (e.g. one can work in more than one country).

In the past two decades, there has been a growing interest in population changes, including population flows and their implications (e.g. Bell et al., 2002; Andersen and Dalgaard, 2006; Kupiszewski and Kupiszewska, 2010). These transformations generate many serious macroscale consequences, but both their intensity and their pace are most noticeable at lower spatial levels (Okólski, 2004), where global processes are often amplified by local conditions. Meanwhile, most research tends to focus on the global and national levels, often disregarding regional and local ones. Migration processes, being considerably selective (Kanbur, Rapoport, 2005), also modify the social capital accumulated in small territorial units (Parysek, 1997).

The scientific and methodological foundations of contemporary analyses of the space of flows date back to the mid-20th century, when gravity and potential models started to appear in theoretical research into the interactions between centres (Zipf 1946, Stewart 1948), and input-output analyses started to appear in empirical studies (Isard 1951, 1965; Leontief 1956, Ullman 1957). At the time, most research addressed links between US states and later also between European countries (Van der Linden and Oosterhaven, 1995), or between the administrative units of other countries (Boschma and Iammarino 2009; Thissen, Oort and Diodato 2013; López-Bazo and Motellón, 2018; in Poland Chojnicki, 1961). In time, the scientific material on linkages started to build up, with researchers growing aware of the mutual effects of various linkages between territorial units (e.g. Kritz and Zlotnik, 1992). In 1994, Tapinos drew attention, inter alia, to the growing dependence between trade and migration. Current knowledge is robust, but most of it has a global or else strictly local dimension. Global studies address links at the level of countries or groups of countries (macroeconomics). Others relate more to enterprises or industrial sectors (e.g. Timmer at al., 2015). Only some of them are cross-territorial in nature, covering various categories of linkages and flows. These include ESPON research under the 3.4.1, 1.4.3, and FOCI projects. By contrast, sometimes local studies describe in great detail the complexity of linkages between a specific region, city (usually a large metropolis, like London; cf. Reades, Smith, 2014), or even a village (Yingnan et al., 2020; Bush, Oosterveer, 2007) and the environment (also the global one; so-called glocalisation). Another group of studies focuses on international linkages within a limited space, particularly sensitive border regions (Blatter, 2004). The relationship between the space of places and space of flows is also addressed by certain transport-related studies (e.g. regarding air linkages between specific cities; Derruder, Witlox, 2005) and transport corridors (Albrechts, Coppens, 2003). Also, the latest research on linkages and flows is limited to a single category, usually foreign trade or migration. The former mainly includes country-level analyses or internal matrix-based research of the input-output type conducted for the US (Stumpner, 2019) and increasingly also for China (Wang, Li 2019).

1.2 Research need

Despite strong theoretical foundations, the “space of flows” still remains less studied than the “space of places”. This is due to the deficit of public statistics (especially international ones), the need to rely on very extensive matrix data, and the lack of general knowledge about the socio-economic determinants and effects of different types of flows and the mutual effects of different types of spatial interactions between territorial units.

Studies of migration flows have to date concentrated primarily on the issue of migration inflow into the EU from non-European countries (e.g. Sirkeci et. al., 2012; Fassman and İçduygu, 2013; Maddaloni and Moffa, 2019) or on post-accession East-West migrations in the framework of the EU (Kahanec et al., 2009; Barrell et al., 2013; Glorius et al., 2013). These studies are limited to the analysis of flows and their consequences at the national level. On the other hand, investigations relative to lower spatial levels are carried out usually from the perspective of the “space of places” (e.g. distribution of Poles and Lithuanians in the United Kingdom based on WRS data; Grabowska-Lusinska, 2013).

There exists, therefore, a need to examine migrations in Europe not only from the national perspective, but also from the regional one, and in terms not only of the “space of places”, but primarily of the very flows and mutual interregional connections.

In light of all of the above, the following gaps in existing knowledge, which justify the reported study, are particularly relevant:

- There is a lack of unified migration statistics in Europe, on both national and international migration flows. The few studies that exist concern countries, not regions. Numerous inconsistencies are observed between the data on the very same flows, registered in individual countries. Altogether, the issue requires a new methodological approach, to establish a complete migration flow matrix in Europe at the regional level. Elaboration of the new approach may constitute both the recommendations for the European and the national statistical institutions (e.g. on the structure of data collection and the spatial level), and the proposal for a unified methodology of estimating the respective missing data.
- Migration flows ought to be analysed to a greater extent in the context of the socio-economic situation of regions, not entire countries. Traditional push and pull factors do not explain adequately the complexity of the interactions or their directions. It is also very important to analyse simultaneously and with methodological consistency the various kinds of spatial interactions.
- Migration flows (like other kinds of spatial interactions) are the key conditioners of development. Knowledge of them is crucial for the development of effective regional and spatial policies at different scales.

1.3 Objective

Our fundamental objective was to identify the magnitudes and the directions of migration flows inside the ESPON space at two levels: (1) national and (2) regional. The flows were analysed both internationally (external migrations) and domestically (internal migrations) for the period of 2010-2018.

This was an enormous challenge because of the significant dispersion of data, various methodologies of gathering these data, incomplete time series, various definitions of migrants, and so on. Our methodology (see Chapter 2) enabled us to construct a matrix of flows between all the ESPON space countries (32 x 32 countries) and between regions in the countries within the ESPON space (297 x 297 regions).

The objectives of the study also included: identification of factors influencing the magnitude of migration flows (explanatory factors), recognition of new spatial structures and interdependences, and regional typology, encompassing various aspects of migration flows (intensity, balance, concentration, distance, selectivity).

The study therefore fulfils the following three complementary objectives:

- Methodologically we have proposed a consistent method by which to estimate a complete migration matrix at the regional level, elaborated a set of universal indicators, reflecting the content of the migration matrix, and tested the econometric models, enabling an explanation of the spatial distribution of migration interactions.
- Empirically, we have recognised the system of migration connections in Europe at the national and regional levels dynamically (2010-2018), assessed the respective balances, concentrations, and associations with geographical distance, and identified the explanatory factors.
- Policy-wise, we have made it possible, with the help of our matrices, to assess the position of each of region within the network of migration flows. This is the key component for the development of a regional policy in the various countries concerned and of a cohesion policy at the level of the European Union.

2 Methodology

2.1 Migration data harmonization strategy

This section provides high-level outlook of strategy as applied to the harmonization of both C2C and R2R flows. To start with, in approaching migration data we faced the well-known problem of differences between migrant counts reported by origin and destination country. The problem exists even when migrant counts adhere to the common migrant definition under the EUROSTAT umbrella. An optimal way to proceed would be to get deep insight into national statistics, data gathering methods, and sources of discrepancies, and then seek for their systematic structure over EU countries. Both solutions impose significant methodological challenges and time constraints and are simply not feasible within the scope of this project. Instead, we had to clearly acknowledge several constraints and incorporate them into our methodology.

1. There is no way to give preference to either origin-reporting or destination-reporting for all the countries concerned. This precludes a global decision.
2. There is no unbiased, scientifically sound way of discerning 'better' and 'worse' statistics at the level of individual countries. Even if we managed it, this kind of evaluation would be difficult to operationalize. This precludes individualized approach to country pairs.
3. Migration flows are most often underestimated. This basically comes from:
   a. the delay and dispersal of information among institutions of a single country (with critical threshold of one year, constituting migrant status, evading administrative procedures and monitoring);
   b. lack of co-ordination between countries;
   c. lack of incentive among migrants themselves to reveal their status for psychological, social, and economic reasons.

Without going into further discussion, the above constraints led us to a single operational principle: **aggressive counting** — i.e. extracting the highest number of migrants wherever sources are contradictory and leaning towards positive errors (surplus of migrants) wherever estimates allowed a certain flexibility. By no means perfect, this attitude at least allows our errors to compensate for national statistics loopholes rather than reinforce them.

Another principle, stemming from sparseness of data, was to accept **data heterogeneity**. Especially on further stages of processing, we had to run estimations on datasets composed of elements from multiple sources, having various reliability and sometimes also coming from earlier estimates.

As our workflow developed many internal data dependencies, we also decided to make our procedure **strictly layered**, in that we treated data obtained or estimated at earlier stage as fixed and did not accept backtracking and improvements based on data from later stages. By this rule, we had to deal with a tree of dependencies instead of a uncontrolled network of dependencies. We hope this approach will also contribute to better understanding of the data by the wider public.

The paramount goal of harmonization – exhaustive coverage of flows at NUTS2 level – was constantly conflicting with the lack of data. This concerned not only true migration data, but also other related datasets, which could be helpful for estimation: they were either not available, had too low resolution / incomplete time series or were the result of crude estimations. However, our goal was not to develop comprehensive modelling framework, but to harmonize whatever data was available. Therefore, we decided for strategy of patching the gaps with first – the most relevant or reliable data, and then – with weaker data, finally arriving at 'last resort' data. By weaker data, we considered aggregates of outflow (at origin side) and existing migrants stock (at destination side). At the point of last resort, we made an assumption that:

- pushing force for migration (outflow driver) is population volume in sending spatial unit,
- attracting force (inflow driver) for receiving spatial unit is GDP (Gross Domestic Product).

We believe, when none of better proxies are not available, sheer population volume is a pool of prospective migrants among whom a certain number will indeed decide to change the residence and seek for a better future. On the other hand, economic activity is most encompassing proxy of attractiveness, with internal links to employment, wealth and quality of life sought by migrants. We are aware these are too bold and undiscriminating factors for so complex and spatially diverse phenomenon like migration but, under circumstances – the only ones plausible. It has to be noted that secondary data like population and GDP were used mostly for approximating the spatial distribution of flows and not their absolute magnitude.

The patching strategy resulted in a maze of data blocks (groups or individual C2C or R2R flow values) with different origin. The pattern changed year by year. Except for stating that 'data had varying origin', it was impossible to curb it into systematic and clear description. On the other hand, we believe that understanding data sources and transformations is important for ourselves, IRiE team members and finally – end users. This requirement is especially true when part of output data is of questionable quality. Due to complexity, is was not sufficient to provide overall description (for most of data items) and then cover special cases with footnotes, as it often happens with statistical sources. The other disadvantage of this widely used scheme is its inability to categorize exceptions into wider classes and inability to arrange the data items along some quality axis.

To handle these problems, we introduced a system of **data provenances**, which consists of

- a dictionary of codes describing data origin (source institution or transformation),
- additional datasets mirroring numeric data.

The system covers every data table used as an input for computations and propagates through the processing up to final output tables. The system can work two ways: 1) as a planning tool to guide data item creation/processing in teamwork setting and 2) as a documentation tool after data items have been created. Data provenance code is assigned to each cell of data individually, so in effect, every data structure processed has a mirror structure for provenience with the same number and layout of cells. Any meaningful difference of data provenance is registered, especially influencing data credibility[1]. So for instance, for a background composed of mostly raw unprocessed data, one year block was missing and patched with averages of nearby cells, the relevant block of provenance cells would be coded "avg", as opposed to "r" for the rest. Codes dictionaries are local to single dataset, because, as a rule, they describe specific issues, relevant to single dataset. When two or more data tables are combined to obtain processed data, their provenance codes are combined too. De-coding of such a provenance must be done by inspecting source tables' dictionaries.

Additional goal of provenances system was to establish a **grading** of data sources. However, fully objective grading is unrealistic and the only feasible approach was to establish grading relative to 1) our trust in data providers and 2) our knowledge of data transformations and their 'severity'. We did attempt to develop ‘grading algebra’, because any score based on interval or ratio scale would require extensive (and probably – controversial) theoretical framework. Instead, provenances are listed in decreasing grading order (best first). Users are encouraged to consult provenances system, make own judgements on data credibility and use them in unbiased scientific reasoning and honest presentation. General concept of proveniences can accomplished with varying degree of formalism and completeness. Improvements in handling data transformations and in grading are definitely possible[2], but for the task at hand we did not attempt to make it more sophisticated then necessary.

2.2 Data creation: Country to Country matrixes

At C2C level, harmonization procedure consisted of three steps consecutively filling the flow matrix up to full content. These were:

1) Base Data,
2) Stock Gain estimation,
3) In-Out-Cross estimation.

2.2.1 Base Data

Primary source for Country-to-country migration matrix was of EUROSTAT/national statistics migrant counts reported either by origin or destination country. Data was collected from two sources: the EUROSTAT database (primary source) and statistical offices (NSI) of all countries of the ESPON space (secondary source). EUROSTAT database contained full data (flows in

[1] We resigned from differentiating between NSI (National Statistical Institution) and Eurostat as separate data provenances, because they essentially refer to the same data passed along.

[2] If required, order-scale multiplication can be applied, because 3 data sources are numerically tied by multiplication too.

both directions) for 18 countries belonging to ESPON space, and incomplete (only for selected countries) for 2 countries belonging to ESPON space. In the first stage, statistical data was collected from publicly available NSI websites, second stage involved enquiries about the missing data. These were send to all countries from ESPON space. Almost all NSIs responded to this enquiries (except from Lithuania, Ireland and Belgium)[3]. However, in the majority of cases, the NSIs' responses confirmed the prior information obtained from the preliminary research that there was a shortage of data (in particular, at the NUTS 2 level). In few cases only, the obtained data turned out to be an important addition to the previously collected data (Hungary). Raw data obtained from EUROSTAT and NSI left us with the following problems:

- lack of data in one direction (outflow or inflow),
- incomplete time series,
- incomplete matrix of flows (e.g., only main migration directions - Top 5, Top 10),
- data aggregation preventing identification of particular countries (e.g., EU-27, non-EU, developed countries / developing countries).

Raw data was re-arranged to C2C 32×32 matrix format with origin countries running row-wise and destination countries running column-wise, which involved transposition in case of destination-reported data. This resulted in 2 alternative sets of 10 single-year matrices of 992 cells each. The completeness of data was:

- from 600 (for 2010) to 685 (2018) valid data cells for origin-reported data,
- from 639 to 680 cells for destination-reported data.

Either way, raw data could only provide no more than ⅔ of total number of cells required. The next step was then to combine both sources in order to take advantage of possible non-overlap of both data sets. According to principle of aggressive counting, the formula of combination was to pick non-empty value if one of two was missing or the higher of two values if both were available.

The number of valid cells resulting from the combination of origin-reported and destination-reported counts was from 852 (year 2010) to 900 (2018). The considerable raise was due to the fact that migrants from countries reluctant to monitor their outflow were often subject of inflow monitoring in more migrant-savvy destination country. It seems to be a rule that destination country is better motivated to track incoming migrants. This phenomenon was particularly visible for Switzerland, where – for the most of time series – inflow counts for as much as 27-28 countries outperformed those reported by senders, as well as Netherlands with 25-27, closely followed by Belgium, Denmark, Germany, Ireland, Luxembourg, and Austria.

Resulting single set of 10 matrices was filled with roughly half of origin-reported and half of destination-reported cells.

[3] With United Kingdom response restricted to a pricelist only.

Except for eliminating empty cells, the procedure revealed true scale of data discrepancies between origin and destination-based reporting. It was quite common to observe both values differ by factor of 2, occasionally reaching above 20. Single exceptionally high discrepancies could reach factor of 217 between Germany and Romania and 192 was between Bulgaria and Slovenia in 2018. Per country, highest discrepancies were observed in outflows from Slovakia, Bulgaria, Romania and Poland[4]. However, these could only be revealed for countries with two-way reporting. Countries with significant number of single-side reported cells, like United Kingdom, Czech Republic, Latvia, Portugal may very well be suspected of errors in migrant counts, but verification was not possible.

Table 2.1: Provenances resulting from first stage were as follow

Code	Grading	Short Name	Applied when	Formula
r	0	Raw data (reported by origin country)	t = null	x = r
t	0	Transposed data (reported by destination country)	r = null	x = t
r<f>	-1	Raw wins	t ≠ null r ≠ null r ≥ t	x = max (r, t) f = int (x / min (r, t)) ratio of evaluated value to discarded value (always ≥ 1) int() is integer rounding function
t<f>	-1	Transposed wins	t ≠ null r ≠ null r < t	

Two variants of ‘r’ and ‘t’ code distinguish data derived from single-side and two-way reporting. The grading of single-side case is higher, because we have no evidence of discrepancy.

As an example of actual provenances, a code for Bulgaria-Belgium flow for 2018 was 't2' meaning the count source was destination country (Belgium) and that origin-based count also existed, but was twice as low (rounded to nearest integer value).

2.2.2 Stock-based estimation

Next set of C2C cells was estimated based on migrants stock (permanent inhabitants of foreign origin). Here, an assumption was that new migrants follow established spatial migration pattern. There are two strong arguments to support this assumption, namely:

[4] Not counting factor of "infinity", which was also present when one of sources reported 0 count.

- family and relatives provide good support for new wave of migrants, facilitating everyday maintenance, job acquisition, and cultural adaptation at destination site;
- information on destination site spreads at origin country and stimulates migration decision.

We did not attempt to explore stock changes year-to-year, and perform any kind od demographic balancing, because this would require full-blown demographic model with high data demands. Used in our straightaway manner, migrants stock is a conservative proxy of distribution, preserving current preferences and not seeking for trendy new directions.

Raw data source was EUROSTAT "Population by country of birth" table. For further calculations, migrants **stock** $s_{year}[x, y]$ was the number of x-born inhabitants of destination country y, reported on January 1, year+1. Three dimensional matrix was first "flattened" (broken into 2-dimensional sheets) by year and then referenced row-wise by country of origin and column-wise by country of inhabitancy. Notation s[y] is used in context of fixed, single year and fixed, single destination country. Notation s_{year} is used in context of fixed pair of destination and origin countries.

A number of C2C flow cells estimated with stock method was in range of 42 (2018) to 75 (2010), mostly occurring in inflow to Czech Republic, Ireland, Latvia, Romania, United Kingdom, as well as Bulgaria until 2011 and France until 2012.

2.2.3 Stock harmonization steps

Before use in C2C estimation, stock table had to be harmonized with several steps.

- For destination countries with partial time series having at least 2 valid entries (mostly ≥7, mostly newest) missing entries were derived from valid entries by means of the average year-to-year change rate for known years.

For each destination country, a series of change ratios r_i was computed for all pairs of non-null stocks for consecutive years t and t+1:

$$r_i = \frac{s_t}{s_{t+1}}.$$

For m non-null stocks, the number of ratios is $n=m-1$. Average change rate d is then harmonic mean

$$d = \frac{n}{\sum_{i=1}^{n} \frac{1}{r_i}}.$$

Missing values occurred in contiguous blocks either at the start or the end of the 2010-2019 series. Estimation was based on closest known year. Estimated stock x for missing year t and known value for subsequent tear x_{t+1} (backward case) was then

$$x_t = x_{t+1} d.$$

In opposite direction (forward case), *x* for missing year *t* and known stock for previous tear x_{t-1} was

$$x_t = x_{t-1}d.$$

After calculation of initial estimated value, the formula was repeated for the rest of missing entries in a block, moving backwards or forwards.

- Ireland-born migrants stock was only available for 2010 and assumed constant over 2011-19.
- For the rest of the cases, missing countries remained constantly missing over time and there was no supplementary data for derivation. These were left empty and excluded from further processing.

At the exit of harmonization, stock data had either complete time coverage for a country pair or no coverage at all.

A set of non-empty stock values for particular destination country (a row) defined a stock-set. In further computations, this set was a base for estimating flow by stock-gain method. Although stock data had their provenances, we decided not to make them part of general provenance system for this would unnecessarily complicate final codes for R2R metadata.

2.2.4 Stock Gain method

In stock gain method, we targeted some of the missing flow values by establishing a model between stock volume *s* (explanatory variable) and flow *f*, namely:

$$f = a + bs.$$

A separate model existed for each origin country *x* for a single year. The model had support in those destination countries *y*, for which both stock and flow is known and later applied to those countries *y*, for which flow is unknown. Both sets varied from country to country and some formalism must have been introduced to handle this. Regardless of their origin, non-empty flow values for destination country *x* (a row in flow matrix) constitute set V defined as follows:

$$\mathrm{V}: \{y \mid flow(x, y) \neq null\}.$$

Also, for the same country *x*, non-empty entries in stock matrix constitute set S:

$$\mathrm{S}: \{y \mid stock(x, y) \neq null\}.$$

Based on above, we distinguished two sets of destination countries with different roles. First was a model support set:

$$\mathrm{I} = \mathrm{V} \cap \mathrm{S}.$$

This set provided data for building a relationship between stock and flow. Second was model result set, composed of countries with unknown flows and known stocks:

$$\mathrm{R} = \mathrm{S} \setminus \mathrm{I}.$$

This set covered estimated flow values for missing countries. After I and R sets were established for destination country, a linear model could be re-written more specifically as

$$f_{\in R} = a + b s_{\in I}$$

and then used for estimation of unknown values.

The actual size of model support set I was never smaller than 16 cases in 2010 and 18 cases in 2011-2018. The actual value of R^2 for linear model can be inspected in provenance code.

Table 2.2: Provenance code added after stock-based estimation was

Code	Grading	Short Name	Applied when	Formula
Stock (<R2>)	-2	Stock Gain	cell in S \ I set	S: {y | s(x,y)≠null} a set of countries with valid stock of x origin T: {y | t(x,y)≠null} a set of countries with valid flow from x model support set I = T ∩ S a, b = parameters of linear model y = at + b over I result set R = S\I x = at + b over R

As an example, the flow between Ireland and Czech Republic in 2018 was assigned code 'Stock (0.905)' to indicate that linear regression was fitted with 0.905 R^2. It can be noticed that the same model was used for Ireland-Latvia and Ireland-Romania (the same row in C2C matrix).

2.2.5 In-Out-Cross

After applying Stock-Gain method, missing cells still existed for the following countries: CZ, IE, EL, ES, CV, LV, MT, PL, PT, RO, UK. For these countries, we established two alternative linear models:

- Outflow model based on population as explanatory variable,
- Inflow model based on GDP as explanatory variable.

C2C flow matrix row was a source of dependent variable for outflow model. For a given country of interest, data for outflow model was derived from a respective row and data for inflow – from a respective column. C2C matrix was already filled with Base data and Stock-estimated data, so size of support set was bigger then in previous step, however we had to accept heterogeneity.

The procedure was similar to Stock-gain method, except that I and R sets with their data (dependent variable) had to be fetched twice – once for column and once for row.

At the final step, both models were compared in terms of explanatory strength of underlying model (R^2) and the winning model was used to generate predicted value. The decision was made based on average R^2 for entire time series, so it is guaranteed that In-Out-Cross cells for a single country are derived from the same type of the model (either Pop or GDP). Otherwise, year-by-year uncontrolled flipping of estimated values could be observed. On rare occasions (approximately 9 cells per year), chosen model predicted negative migrant counts. This happened for some inflows/outflows to/from Malta and Cyprus – apparently due to their lowest absolute migrant counts. These values have been artificially upgraded to 0 and marked in provenance code.

Table 2.3: Provenance codes added after In-Out-Cross estimation were:

Code	Grading	Short Name	Applied when	Formula
Pop (<R2>)	-3	Population-based Outflow	Pop R^2 (2010-18) > GDP R^2 (2010-18)	see Stock-Gain
GDP (<R2>)	-3	GDP-based Inflow	GDP R^2 (2010-18) > Pop R^2 (2010-18)	
Pop SubZ (<R2>)	-4	Population-based Outflow	as above but estimation upgraded to zero	
GDP SubZ (<R2>)	-4	GDP-based Inflow		

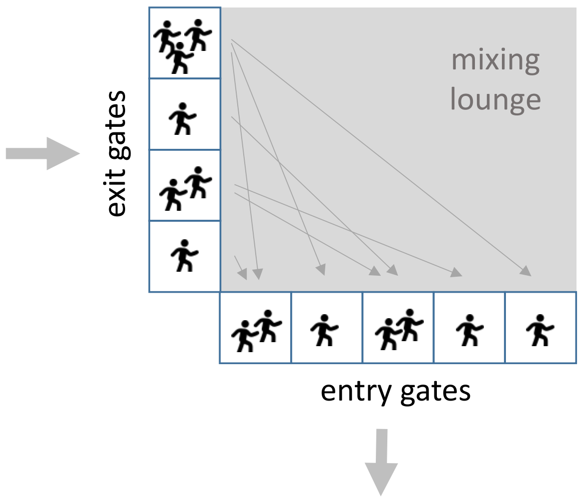
As an example, missing inflows to Portugal 2018 were estimated with outflow Pop-based model (from Ireland, Cyprus and Malta) and inflow GDP-based model (from Czech Rep., Greece, Latvia, Poland, Romania and United Kingdom). The same Pop-based model was used for all cells, with R^2=0.387. On the other hand, each cell estimated from GDP came from different inflow model, with R^2 ranging from 0.469 to 0.6.

2.3 Data creation: Region to Region matrixes

Our thinking about estimation of R2R flows was largely determined by existence of C2C flows. This diverted us from building of any kind of independent regional model(s) based on demographic or economic factors. The task at hand was to distribute a known migrant counts over the spatial units, rather than estimate migrant counts from ground up, using population dynamics and structure inherent to each spatial unit separately. We also sought for safe and straightforward method, especially because model verification based on observed data was not possible. Thus, we introduced a **double-gates** model, predicting the distribution of migrants as a combination of two forces: pushing and pulling, referred as outflow driver and inflow driver. We also made a very strong conjecture that these forces are **independent**. This conjecture was necessary because no data on regional preferences of migration exists, so single 'combination operator' must apply to whole structure. If this was not the case, then some kind of proxy data of R2R structure would be necessary. One of such sources would be distance/cost of travel between regions, leading to gravity model. This was abandoned early as lacking rationale, especially in EU conditions[5].

From the perspective of decision process of a single prospective migrant, independence conjecture translates to disjoint, two-step process: one decision to leave the country of origin and second decision to pick a specific place in destination country.

Figure 2.1: Double-gates migration model



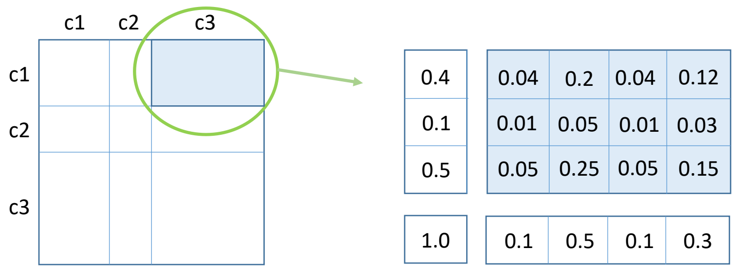
As depicted on **¡Error! No se encuentra el origen de la referencia.**, a flow of imaginary migrants leaves origin country and reports at a row of exit gates, according to their region (NUTS 2) of origin. Then, migrants pass to a mixing lounge – a void space where they can freely make decisions about their destination in receiving country. Finally, migrants report at a row of entry gates, disclosing their preference to the public.

The whole process is confined to a single origin-destination country pair and so is the numerical procedure. R2R distribution matrix (with rows as origin regions and columns as destination regions) is split into rectangular blocks of cells belonging to single origin country and single desti-

[5] However gravity hypothesis is tested in separate subtask of the project.

nation country (see **¡Error! No se encuentra el origen de la referencia.**.2). Whatever follows is performed in one of 32×32–32 = 992 blocks.

Figure 2.2: R-R Matrix breakdown



Within the block, exit gates are approximated by outflow driver vector (left margin) and entry gates are approximated by inflow driver vector (bottom margin). Both drivers are expressed as shares, summing up to 1. Now independence condition allows us to calculate joint probability as a simple product of two independent probabilities. The result is the share of migrant flow falling to single region-to-region pair, summing up to 1 for country-to-country pair.

For outflow share $Out_{o,d}$, inflow share $In_{o,d}$ and country-to-country flow of migrants *C*, the flow *T* between origin and destination regions *o,d* is

$$T_{o,d} = Out_{o,d} * In_{o,d} * C.$$

Example calculation of 2010 flow between DK01 and CZ01 units.

- Outflow distribution for Denmark assigns 0.472 of outbound migrants of Denmark to DK01 unit for year 2010 (distribution varies year by year);
- Inflow distribution for Bohemia, specific for Danes, assigns 0.614 of total Denmark inflow to CZ01 unit.

Share of DK01 migrants arriving in CZ01 is then 0.29, out of 1.0 for all DK units and all CZ units.

- Since C2C flow between DK and CZ is 443 persons, DK01→CZ01 flow is 128.4 persons.

2.4 R2R data sources

R2R harmonization guideline idea can be summarized as follows: *Migrants adhere to country-wide pattern at source and imitate previous migrants at destination*. We believe this guideline sets down relatively safe and conservative approach, which is best fitted for incomplete and risky data sources.

An input to region-to-region flows estimation consisted of three data structures:

- C2C Migration Flows Matrix (generated as described in section **¡Error! No se encuentra el origen de la referencia.**,
- R-Outflow table of Regional Migration Outflow as Outflow Driver
- R-Stock table of Regional Distribution of foreign population (Stock) as Inflow Driver

Outflow driver was readily available in form of EUROSTAT Migratory Outflow with country resolution and year-by-year coverage. We picked external migration.

Inflow sums were not available directly and had to be approximated. Our working assumption was that *migrants follow established migration pattern*, so the proxy was existing migrants stock, with rationale explained earlier in Stock Gain method for C2C. As a bonus for R2R level, we were able to collect more specific stock data, namely divided into origin country of resident migrant population. This way, preferences of new migrants are related to their country of origin. The downside is that this structure is constant over time series.

Two former data sources were not perfect and had to be harmonized.

2.4.1.1 R-Outflow table harmonization

External and Internal Migratory Outflow data available from EUROSTAT covers BG, CZ, ES, HR, IT, LT, HU, NL, AT, PL, RO, SI, SK, FI, SE, NO, CH. Countries not covered by either external nor internal outflow include BE, DK, DE, IE, EL, FR, PT, UK. Mono-regional countries, for which the structure is irrelevant include EE, CY, LV, LU, MT, IS, LI. Two methods were used to complement the missing data:

1. Extending time series in case of gaps for singular years, namely: Croatia 2010, Italy 2010, Poland 2015, Switzerland 2010. This was performed with harmonic mean, like with -> Stock harmonization steps.
2. Disaggregating the C2C emigration volume to regions by a share of region's total population in total given country's population. This method was applied to Belgium, Germany, Greece, France, Portugal and United Kingdom. Given R2R procedure, this effectively equals using outright population volume as outflow driver, but was performed for the sake of completeness. For Ireland, 1 out of 3 regions was available as raw data and two missing regions were estimated by population share.

Table 2.4: Resulting provenance codes were:

Code	Grading	Short Name	Applied when
r	0	raw data	
TSE	-1	Time Series Extension	raw=null for single year
POP	-2	Population-based estimation	raw=null for all years

2.4.1.2 R-Stock table harmonization

R-Stock table content was foreign population by NUTS2 region (rows), by country of birth (columns). The data come from two sources: (1) NSI (primary source of information); (2) EUROSTAT (only for mono-regional countries – irrelevant for R2R procedure).

Figure 2.3: A snippet from R-Stock table

		BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	HR	IT	CY	LV
Région de Bruxelles-Ca	BE10	766744	5158.102	334.5796	83.6449	6245.486	139.4082	223.0531	4461.061	6831	17788.48	1003.739	103078.4	27.88163	278.8163
Prov. Antwerpen	BE21	1602721	2551.247	165.4863	41.37157	3089.078	68.95262	110.3242	2206.484	3378.679	8798.355	496.4589	50983.57	13.79052	137.9052
Prov. Limburg (BE)	BE22	766260	1831.878	118.8245	29.70612	2218.057	49.5102	79.21633	1584.327	2426	6317.502	356.4735	36607.84	9.902041	99.02041
Prov. Oost-Vlaanderen	BE23	1380539	1058.437	68.65539	17.16385	1281.567	28.60641	45.77026	915.4052	1401.714	3650.178	205.9662	21151.58	5.721283	57.21283
Prov. Vlaams-Brabant	BE24	999933	1475.173	95.68688	23.92172	1786.155	39.86953	63.79125	1275.825	1953.607	5087.352	287.0606	29479.53	7.973907	79.73907
Prov. West-Vlaanderer	BE25	1129199	809.766	52.52536	13.13134	980.4735	21.88557	35.01691	700.3382	1072.393	2792.599	157.5761	16182.19	4.377114	43.77114
Prov. Brabant wallon	BE31	348905	723.6305	46.93819	11.73455	876.1796	19.55758	31.29213	625.8426	958.3214	2495.547	140.8146	14460.87	3.911516	39.11516
Prov. Hainaut	BE32	1166649	3328.058	215.8741	53.96851	4029.649	89.94752	143.916	2878.321	4407.429	11477.3	647.6222	66507.2	17.9895	179.895
Prov. Liège	BE33	961736	2509.312	162.7662	40.69155	3038.302	67.81924	108.5108	2170.216	3323.143	8653.735	488.2985	50145.55	13.56385	135.6385
Prov. Luxembourg (BE)	BE34	254296	632.0204	40.99592	10.24898	765.2571	17.08163	27.33061	546.6122	837	2179.616	122.9878	12630.16	3.416327	34.16327
Prov. Namur	BE35	455028	566.9468	36.77493	9.193732	686.4653	15.32289	24.51662	490.3324	750.8214	1955.2	110.3248	11329.74	3.064577	30.64577
Severozapaden	BG31	8.478941	530920.3	26.95869	3.40607	61.45421	2.246557	5.652628	90.80439	11.23279	20.14654	2.753844	33.04613	8.044124	4.710523
Severen tsentralen	BG32	9.224932	577631.5	29.33055	3.705742	66.86104	2.444213	6.149955	98.7935	12.22106	21.91907	2.996132	35.95358	8.751858	5.124962
Severoiztochen	BG33	12.51474	783627.3	39.79046	5.027289	90.70513	3.315872	8.343161	134.0254	16.57936	29.73588	4.064617	48.7754	11.87296	6.952634
Yugoiztochen	BG34	13.93196	872368.5	44.29649	5.5966	100.977	3.691374	9.287974	149.203	18.45687	33.10329	4.524911	54.29893	13.2175	7.739979
Yugozapaden	BG41	56.44529	3534397	179.4671	22.6746	409.1077	14.95559	37.63019	604.4953	74.77794	134.1179	18.33266	219.9919	53.55066	31.35849
Yuzhen tsentralen	BG42	16.40414	1027167	52.15675	6.589696	118.8949	4.346395	10.93609	175.6785	21.73198	38.97735	5.32784	63.93408	15.5629	9.11341
Praha	CZ01	255	2629	1080546	191	3120	64	483	280	486	1839	1305	1558	62	109
Strední Cechy	CZ02	83	1284	1223015	38	1008	11	45	62	42	194	185	210	10	40
Jihozápad	CZ03	30	807	1149939	2	2803	3	13	28	30	127	145	180	6	16
Severozápad	CZ04	38	651	1050596	32	6145	8	12	38	36	89	268	186	3	26
Severovýchod	CZ05	27	763	1441092	13	1010	4	19	68	29	95	117	149	8	12
Jihovýchod	CZ06	45	686	1614853	18	740	12	26	142	67	207	247	278	2	22
Strední Morava	CZ07	22	216	1186636	9	596	2	19	99	31	75	95	218	3	10
Moravskoslezsko	CZ08	23	399	1177367	8	341	7	8	222	30	89	125	154	1	8
Hovedstaden	DK01	438	1776	406	1547253	6896	475	803	764	2366	3626	324	3273	26	875
Sjælland	DK02	90	271	62	783752	1551	70	99	65	240	256	45	266	4	286
Syddanmark	DK03	115	380	151	1136485	8038	146	200	107	381	535	58	486	4	1091
Midtjylland	DK04	195	1041	173	1198372	4162	320	253	158	742	675	62	712	5	1091
Nordjylland	DK05	67	335	57	555042	1421	129	58	50	195	329	21	261	2	409

Three methods were used to complement the missing data:

1. Disaggregation of data on foreign population at the national level by breaking data down into the regional level according to region's share of GDP (euro) in total given country's GDP (e.g., France, UK);
2. Disaggregation of data on population by breaking these data down into subtotals by citizenship at the regional level (Belgium);
3. Disaggregation of data based on countries' GDP was used in the absence of information on foreign population of selected nationality (World Bank estimates). Given R2R procedure, this effectively equals using outright GDP as inflow driver, but was performed for completeness.

Table 2.5: Resulting provenance codes were:

Code	Grading	Short Name	Applied when
r	0	raw data	

GDP1	-1	GDP-based variant 1	estimation	raw=null
GDP2	-1	GDP-based variant 2	estimation	raw=null, regional sums available

2.4.2 Final Provenances for R2R flows

Compound provenances included 'Migration R2R Meta.xlsx' were computed by combining source provenances of 3 sources (C2C, R-Outflow and R-Stock), according to the same pattern as in numerical calculation. Simple text concatenation was used, with long dash character was used as first separator and short dash (minus) character as second separator. Total grading of particular R2R cell must be assessed manually by combining original grades.

2.5 Data creation: Region to Region matrixes (domestic migration)

In the case of data on the internal interregional flows (i.e. taking place inside one country) only for five countries (Bulgaria, Spain, Italy, Austria and Norway) the complete data sets for the years 2010-2018 were at disposal. For the remaining countries gaps existed, associated, in particular, with incomplete time series, or aggregation of data to other spatial levels than NUTS 2. Information on the availability of data in the particular countries and on the undertakings aimed at estimation of domestic flows is provided in Table 2.6.

Table 2.6: Internal migration – data availability

Country	Data sources	Description
Belgium	Statistics Flanders IWEPS	Matrix of flows between the Flemish NUTS 2 units and NUTS 1 units (2010-2017) available. Data are available on migration inflows and outflows according to the communes in Wallonia as well as the matrix of flows for 2019 in Wallonia. Value of each single cell of internal flow matrix for NUTS 1 (main diagonal excluded) was calculated as an appropriate share within a total disaggregated volume, proportionally to the multiplication of total outflow of origin (raw sum) and total inflow of destination (column sum). Although such approach leads to the situation, that neither raw's nor column's sum is kept with no distortion, it prioritizes a proportion across estimated values within particular cells of the whole matrix instead. Time series extended for 2018 by harmonic mean change[1] over 2010-2017.
Bulgaria	NSI	Raw data at the level of NUTS 2 available for the years 2010-2019
Czechia	NSI	Internal migration matrix according to NUTS 3 (2010-2019) is available; the data were aggregated to NUTS 2 level with omission of flows within the NUTS 2 units
Denmark	NSI	Internal migration matrix according to NUTS 3 (2010-2019) is available; the data were aggregated to NUTS 2 level with omission of flows within the NUTS 2 units
Germany	NSI for the country and the particular lands	Flow matrix for the years 2010-2018 at the NUTS 1 level is available (a part of these units constitutes also NUTS 2 units). For the majority of lands (NUTS 1) data are available at the NUTS 2 level (matrices of intra-land flows and flows with respect to other lands). Flows are available between the NUTS 1 level and all the other NUTS 2 units, re-scaled to the NUTS 2 level by the GDP value or population number (attraction/production), changing yearly. Value of each single cell of internal flow matrix for NUTS 1 (main diagonal excluded) was calculated as an appropriate share within a total disaggregated volume, proportionally to the multiplication of total outflow of origin (raw sum) and total inflow of destination (column sum). Although such approach leads to the situation, that neither raw's nor column's sum is kept with no distortion, it prioritizes a proportion across estimated values within particular cells of the whole matrix instead. In case of data gaps we faced, a procedure of exploiting temporal trends relied on extrapolation forward, as data for 2018 was estimated and earlier data was available only. A geometric row has been constructed, based on the last two available temporal chain's elements has been constructed and extended forward until element of 2018. Taking into account two last available values only allows to avoid chain discontinuity, which would be an effect of undertaken procedure only and be not conditioned by empirical state. Whereas geometric row has been assumed as optimal, because it excludes eventual result of negative value and allow to avoid exceeding constraints of the space of potential data range. On the other hand, an unrealistically dynamic and strongly incriminating results increase of values is possible in such circumstances. It could be derived from either magnification of real observed increase due to geometric rule or specificity of two only temporal intersections taken into account. Therefore, to avoid this risk, one exception from these overriding rules has been accepted and in every single case of increase of values between the last two available temporal chain's elements (where negative values can not occur) a geometric raw has been substituted by arithmetic one.

Country	Data sources	Description
Estonia		The entire country = NUTS 2
Ireland	NSI	Data on internal migrations available only for the years2011 and 2016 according to counties (local government areas); data were aggregated to NUTS 2 level with omission of flows inside NUTS 2 units. Missing 2012-2015 interpolated linearly from 2011 and 2016 datapoints. Then, time series extended for 2010 and 2017-18 by harmonic mean change[1] over 2011-2016.
Greece	NSI	Data on domestic migrations available only for the year 2011 (Census); gaps were filled out with data for 2011
Spain	NSI	Raw data available at NUTS 2 level for 2010-2019
France	NSI	Data available on persons participating in domestic migration flows (change of residence location), 2013-2017; data were aggregated to NUTS 2 with omission of flows between communes, taking place inside NUTS 2. Time series extended for 2010, 2011, 2012 and 2018 by harmonic mean change[1] over 2013-2017. Missing Mayotte inflows estimated afterwards based on La Reunion inflows (Reunion inflow scaled by Mayotte/Reunion population ratio, changing yearly).
Croatia	NSI	Data available on inward and outward movements at NUTS 3 level (2010-2019), followed by the aggregation of data to NUTS 2 level. Value of each single cell of internal flow matrix for NUTS 3 (main diagonal excluded) was calculated as an appropriate share within a total disaggregated volume, proportionally to the multiplication of total outflow of origin (raw sum) and total inflow of destination (column sum). Although such approach leads to the situation, that neither raw's nor column's sum is kept with no distortion, it prioritizes a proportion across estimated values within particular cells of the whole matrix instead.
Italy	NSI	Raw data available at NUTS 2 level for 2010-2018.
Cyprus		The entire country = NUTS 2
Latvia		The entire country = NUTS 2
Lithuania	NSI, Demographic Yearbook	Data on internal migrations available for the years 2011 and 2013-2017. Missing 2012 interpolated as 2011 and 2013 average. Then, time series extended for 2010 and 2018 by harmonic mean change[1] over 2011-2017.
Luxembourg		The entire country = NUTS 2
Hungary	NSI	Raw data available for 2013-2018. Time series extended for 2010, 2011 and 2012 by harmonic mean change[1] over 2013-2018.
Malta		The entire country = NUTS 2
Netherlands	NSI	Available matrix of flows according to communes (2011-2018); data were aggregated to NUTS 2 level with omission of flows between communes, taking place within NUTS 2 units. Time series extended for 2010 by harmonic mean change[1] over 2011-2018.
Austria	NSI	Raw data available at NUTS 2 level, 2010-2019
Poland	NSI	Raw data available at NUTS 2 level, 2010-2019; gaps (i.e. flows between PL91 and PL92) were elaborated from the matrix of flows at LAU 2 level.

Country	Data sources	Description
Portugal	NSI	Available is only the sum of migration flows according to NUTS 2 in the years 1999-2001 and the number of population having changed residence location in 2001 and 2011 (census data). Given lack of other data the average from these years was treated as the population flow in 2001 with consideration of change in time on the basis of data from 2011. The same value was complemented for the successive years.
Romania	NSI	Data available on inward and outward movements at NUTS 3 level (2010-2018) and the matrix of flows at NUTS 3 level (2011). Matrix of domestic flows was used to rescale the annual values of flows, the data aggregated then at NUTS 2 level.
Slovenia	NSI	Matrix of domestic migration flows according to NUTS 3 (2010-2019) is available; data were aggregated to NUTS 2 level with omission of flows inside NUTS 2 units.
Slovakia	NSI	Matrix of domestic migration flows according to NUTS 3 (2010-2019) is available; data were aggregated to NUTS 2 level with omission of flows inside NUTS 2 units.
Finland	NSI	Matrix of domestic migration flows according to NUTS 3 (2010-2018) is available; data were aggregated to NUTS 2 level with omission of of flows inside NUTS 2 units.
Sweden	NSI	Matrix of domestic migration flows according to NUTS 3 (2010-2019) is available; data were aggregated to NUTS 2 level with omission of of flows inside NUTS 2 units.
United Kingdom	Office for National Statistics The Scottish Government	Matrix of flows at the level of English statistical units and the flows between England and Northern Ireland are available for England. Matrix of flows between the Scottish statistical units is available for Scotland. Flows are available between Scotland (NUTS 1) and all other units, rescaled to NUTS 2 level with GDP value or population number (attraction/production), changing yearly. Time series extended for 2010 and 2011 by harmonic mean change[1] over 2012-2018.
Iceland		The entire country = NUTS 2
Liechtenstein		The entire country = NUTS 2
Norway	NSI	Raw data available at NUTS 2 level, 2010-2019.
Switzerland	NSI	Time series extended for 2010 by harmonic mean change[1] over 2011-2018.
[1] For details see Stock Harmonization Steps in external migration section.		

Source: authors' elaboration.

2.6 Measuring migration flows

To measure and describe the migration flows within the research area comprehensively and complementarily, a system of intentionally constructed indexes, ordered by use of overarching four dimensional framework (intensity, balance, concentration and distance impact) has been accepted initially, with full aware of demand for their theoretical independency. They are as follows:

Intensity dimension:

- Intensity index calculated for a pair of regions i and j – relation (i,j):

$$Int_{i,j} = Flow_{i,j} + Flow_{j,i}$$

or calculated for a spatial unit i:

$$Int_i = \sum_{j=1}^{296} Flow_{i,j} + Flow_{j,i}$$

- Weighted intensity index for spatial unit i:

$$Wint_i = \sum_{j=1}^{296} \frac{Flow_{i,j} + Flow_{j,i}}{Pop_i}$$

Balance dimension:

- Balance index for spatial unit i:

$$B_i = \frac{\sum_{j=1}^{296} Flow_{j,i} - Flow_{i,j}}{max\left(\sum_{j=1}^{296} Flow_{i,j}\,; \sum_{j=1}^{296} Flow_{j,i}\right)}$$

- Unbalanced volume index for spatial unit i:

$$BvAsym_i = \sum_{j=1}^{296} Flow_{j,i} - Flow_{i,j}$$

- Average relation asymmetry for spatial unit i:

$$BrAsym_i = \sum_{j=1}^{296} \frac{\left|Flow_{j,i} - Flow_{i,j}\right|}{296 \times max\left(Flow_{i,j}; Flow_{j,i}\right)}$$

Concentration dimension:

- Concentration per area index for spatial unit i, related to total volume of flow:

$$Garea_i = \frac{\sum_{j=1}^{296}\sum_{k=1}^{296} Area_j \times Area_k \times \left|\frac{\left(Flow_{i,j} + Flow_{j,i}\right)}{Area_j} - \frac{\left(Flow_{i,k} + Flow_{k,i}\right)}{Area_k}\right|}{2 \times \left(\sum_{j=1}^{296} Area_j\right)^2 \times \frac{\sum_{j=1}^{296} Area_j \times \left(Flow_{i,j} + Flow_{j,i}\right)}{\sum_{j=1}^{296} Area_j}}$$

to inflow extracted only:

$$Garea_i = \frac{\sum_{j=1}^{296}\sum_{k=1}^{296} Area_j \times Area_k \times \left|\frac{Flow_{j,i}}{Area_j} - \frac{Flow_{k,i}}{Area_k}\right|}{2 \times \left(\sum_{j=1}^{296} Area_j\right)^2 \times \frac{\sum_{j=1}^{296} Area_j \times Flow_{j,i}}{\sum_{j=1}^{296} Area_j}}$$

or to outflow extracted only:

$$Garea_i = \frac{\sum_{j=1}^{296}\sum_{k=1}^{296} Area_j \times Area_k \times \left|\frac{\left(Flow_{i,j}\right)}{Area_j} - \frac{\left(Flow_{i,k}\right)}{Area_k}\right|}{2 \times \left(\sum_{j=1}^{296} Area_j\right)^2 \times \frac{\sum_{j=1}^{296} Area_j \times \left(Flow_{i,j}\right)}{\sum_{j=1}^{296} Area_j}}$$

- Concentration per population index for spatial unit i, related to total volume of flow:

$$Gpop_i = \frac{\sum_{j=1}^{296}\sum_{k=1}^{296} Pop_j \times Pop_k \times \left|\frac{(Flow_{i,j} + Flow_{j,i})}{Pop_j} - \frac{(Flow_{i,k} + Flow_{k,i})}{Pop_k}\right|}{2 \times \left(\sum_{j=1}^{296} Pop_j\right)^2 \times \frac{\sum_{j=1}^{296} Pop_j \times (Flow_{i,j} + Flow_{j,i})}{\sum_{j=1}^{296} Pop_j}}$$

to inflow extracted only:

$$Gpop_i = \frac{\sum_{j=1}^{296}\sum_{k=1}^{296} Pop_j \times Pop_k \times \left|\frac{Flow_{j,i}}{Pop_j} - \frac{Flow_{k,i}}{Pop_k}\right|}{2 \times \left(\sum_{j=1}^{296} Pop_j\right)^2 \times \frac{\sum_{j=1}^{296} Pop_j \times Flow_{j,i}}{\sum_{j=1}^{296} Pop_j}}$$

or to outflow extracted only:

$$Gpop_i = \frac{\sum_{j=1}^{296}\sum_{k=1}^{296} Pop_j \times Pop_k \times \left|\frac{Flow_{i,j}}{Pop_j} - \frac{Flow_{i,k}}{Pop_k}\right|}{2 \times \left(\sum_{j=1}^{296} Pop_j\right)^2 \times \frac{\sum_{j=1}^{296} Pop_j \times Flow_{i,j}}{\sum_{j=1}^{296} Pop_j}}$$

Distance impact dimension:

- Average distance for spatial unit i, related to total volume of flow:

$$Dist_i = \frac{\sum_{j=1}^{296} dist_{i,j} \times (Flow_{i,j} + Flow_{j,i})}{\sum_{j=1}^{296} Flow_{i,j} + Flow_{j,i}}$$

to inflow extracted only:

$$Dist_i = \frac{\sum_{j=1}^{296} dist_{i,j} \times Flow_{j,i}}{\sum_{j=1}^{296} Flow_{j,i}}$$

or to outflow extracted only:

$$Dist_i = \frac{\sum_{j=1}^{296} dist_{i,j} \times Flow_{i,j}}{\sum_{j=1}^{296} Flow_{i,j}}$$

- Distance dependence index for spatial unit i, related to total volume of flow:

$$DistR^2{}_i = \frac{\sum_{j=1}^{296}\left(Flow\ est_{i,j} + Flow\ est_{j,i} - \frac{\sum_{j=1}^{296} Flow_{i,j} + Flow_{j,i}}{296}\right)^2}{\sum_{j=1}^{296}\left(Flow_{i,j} + Flow_{j,i} - \frac{\sum_{j=1}^{296} Flow_{i,j} + Flow_{j,i}}{296}\right)^2}$$

or to inflow extracted only:

$$DistR^2{}_i = \frac{\sum_{j=1}^{296}\left(Flow\ est_{j,i} - \frac{\sum_{j=1}^{296} Flow_{j,i}}{296}\right)^2}{\sum_{j=1}^{296}\left(Flow_{j,i} - \frac{\sum_{j=1}^{296} Flow_{j,i}}{296}\right)^2}$$

or to outflow extracted only:

$$DistR^2{}_i = \frac{\sum_{j=1}^{296}\left(Flow\ est_{i,j} - \frac{\sum_{j=1}^{296} Flow_{i,j}}{296}\right)^2}{\sum_{j=1}^{296}\left(Flow_{i,j} - \frac{\sum_{j=1}^{296} Flow_{i,j}}{296}\right)^2}$$

For the purposes of presentation of balance's dynamics of migration flow at the level of whole R2R matrix, the another index has been developed:

$$B_{total} = \sum_{i=1}^{297}\sum_{j=1}^{297} \frac{\left|Flow_{j,i} - Flow_{i,j}\right| \times \left(Flow_{i,j} + Flow_{j,i}\right)}{max\left(Flow_{i,j}; Flow_{j,i}\right) \times 2 \times \sum_{i=1}^{297}\sum_{j=1}^{297} Flow_{i,j}}$$

The index of intensity is expressing volume of flow within a given relation of for all relations of a given spatial unit in an absolute units of phenomenon (in case of migration - persons). Weighted intensity is a value of the intensity index, related to the population of the given region. It supplements the previous index by the consideration of the context of demographic size of the region. Much smaller volume of flows can have relatively higher importance in case of small region.

Also three indexes of balance are complementary regarding cognitive added value. Balance index indicates if the regions is a net-sender or net-receiver of flow and in how much disproportion between both directions exists. The value range is from -1 to 1, where zero means perfectly balanced flow, -1 means outflow only and 1 means inflow only. The index of balance volume is taking into account the size of a given region's flow and it indicates how much volume of flow is generated or absorbed there. Even largely unbalanced saldo can not generate much surplus or deficit of flow in the whole network if the region is of small size or low weighted intensity. And opposite, regions of the largest agglomerations and of the highest weighted intensity of flow can generate large unbalanced volume of flow even though value of balance index around zero. A complementary role of the average relation asymmetry relies on taking into account balance of each particular relation of the region, even though they can be balanced per saldo. The value range is from 0 to 1, where zero means all of region's relations perfectly balanced and 1 means all region's relations in one direction only. This index can indicate “an intermediate regions”, where inflow and outflow is balanced while regions of inflow and regions of outflow are different.

As balance index calculated for the whole matrix of internal flow within researched area is equal zero by definition, a measure based on balance of each individual relationship instead of it total *per saldo* is more appropriate. Moreover, particular relationships are largely diverse according to intensity index, thus taking into account disparities in their impact on the total spatial structure of migration flow by any kind of weighting is necessary.

Indexes of concentration are using different versions of Gini coefficient, the most commonly used measure of concentration, weighted geographically (by area of particular spatial units) or by population. It is also complementary approach. A concentration per area is indicating pure flow concentration, which is a natural consequence of population distribution inequality to some extent. On the other hand, concentration per population is taking into account this condition and is indicating some kind of concetration's “added value”.

An average distance index express a distance travelled by average migrant of the region, arriving to- or leaving from- the region. It indicates an empirical range of a given region's impact. On the other hand, a distance dependency index indicates a degree of distance impact, regardless of its character (increase, decrease, dynamics etc.). The value of 0 is limiting a range of variability of one side and indicates a lack of any co-occurrence between volume of flow to other regions and distance to them, and 1 is limiting the range of values top and indicates existing function of distance, which describes full variability of flow volume.

An in-depth analysis of four dimensions has been summarized by means of three simple bi-dimensional typologies. They have been developed by classification of one the most representative index for each individual dimension: $Wint_i$ (intensity dimension), B_i (balance dimension), $Garea_i$ for flow total volume (concentration dimension) and $Dist_i$ for flow total volume (dimension of distance impact). Each of these indexes has been classified within the framework of three levels. The thresholds were determined by statistical distribution of empirical data set, where the equal representation is assumed in case of normal distribution (thus $\pm 0{,}431\sigma$ from the mean). Such solution does not imply equal representation of each of three index classes in practice, but returns the added cognitive value of distribution's skewness display. A given pair of classified indexes has been combined furtherly and, finally, each particular combination of classes has been named as appropriate type, by use of terminology expressed at qualitative or/and ordinal scale.

In opposition to individually set of descriptive analysis framework, due to comparability purposes, the set of indexes applied for typology of NUTS 2 according to migration flow is conventional and the same as for each of other flows' individual typology developed under task 1:

- intensity index;
- weighted intensity index;
- connectivity index;
- selectivity index;
- external influence index;
- Send-Receive balance index.

Three indexes listed as the first can be applied either, as related to total volume of flow or only to inflow/outflow extracted.

2.7 Explanatory factors: drivers and barriers

This part of the study was dedicated to identifying the determinants of migration flows between all analyzed regions in Europe. For this purpose, econometric modelling was conducted using panel data considering 297 NUTS 2 regions for the period 2010-2018 where the cross-sectional dimension of the data was determined by all pairs of regions. Hence, the maximum number of observations for each time-varying variable in the model was equal to $N = 87912$ times $T = 9$. It should be noted, however, that for many variables, observations for all regions are not available, resulting in different numbers of observations being used to estimate different specifications of the model explaining migration rates.

The starting point for the econometric model specification was the gravity model, which is used as a basic tool for the analysis of international trade (Leibenstein and Tinbergen, 1966; Baltagi, Egger and Pfaffermayr, 2003), but also migration flows (e.g. Basher and Fachin, 2008; Mayda, 2010; Molloy, Smith and Wozniak, 2011; Beine and Parsons, 2015). Classically, in the gravity model, a key role is assigned to the distance and so called "mass" of both the origin region and the destination region. While in trade analysis mass is represented mainly by total Gross Domestic Product (GDP), in models describing flows of people it is more advisable to take the population size of regions as the mass. Although also many migration studies use the income maximization approach as for example Basher and Fachin (2008), Beine and Parsons (2015), Beine et al. (2019) as well as Serlenga and Shin (2021), there a more important role is attributed to the size of the economy than to the population itself. It seems, however, that in the context of migrations between European regions it is not the size of the country's economy as represented by total GDP that matters more, but rather wealth, which is better reflected by GDP per capita or disposable income per person. Furthermore, it is important to consider that migration decisions in a given year are made not based on knowledge of the level of GDP per capita or income in that same year, but rather on information from previous years. Therefore, lagged values of these variables were used as explanatory factors in the models. It turned out that both the inclusion of one-period and two-period lags gives almost the same estimation results - the tables presenting the results of modeling include estimates of variables lagged by one period.

In addition to the main gravity variables (distance and GDP), the study presented here also considers a number of factors both measurable and non-measurable that can affect migration flows. The number of potential economic, social, geographic and other factors is extremely large, so a preliminary selection was made. The reduction of the number of potential factors was based on the availability of relevant statistical data, but also on their correlation between each other and statistical significance, which was assessed after the initial stages of model estimation.

Hence, the specification of the econometric model used can be written as follows:

$$M_{ijt} = \beta_0 {Pop_{it}}^{\beta_1} {Pop_{jt}}^{\beta_2} {lagGDPpc_{it}}^{\beta_3} {lagGDPpc_{jt}}^{\beta_4} {DIST_{ij}}^{\beta_5} e^{\beta_6 domestic_{ij}} {X_{it}}^{\beta_7} {X_{jt}}^{\beta_8} e^{\beta_9 D_i} e^{\beta_{10} D_j} \quad \mu_{ijt} \varepsilon_{ijt}$$

where: the dependent variable, M_{ijt} refers to the migration flow from a region i (origin) to a region j (destination) in a year t; $GDPpc_{it}$ denotes the GDP per capita of the origin region in period t and $GDPpc_{jt}$ denotes the GDP per capita of the destination region in period t while Pop_{it} denotes the population size of the origin region at period t and Pop_{jt} denotes the population size of the destination region at period t. The geographical distance between the origin region i and the destination region j is represented by $dist_{ij}$ and is measured as the orthodromic distance between the centroids of the regions weighted by the population of the municipalities, corrected for enclaves. Then, the inter-country dummy $domestic_{ij}$ is included, which indicates if the origin region i and the destination region j belong to the same country.

All additional explanatory variables that are expected to affect the volume of migration flows between European regions are represented by X_{it} – time-varying factors for the origin region, X_{jt} – time-varying factors for the destination region, and D_i – time-invariant factors for the origin region and D_j – time-invariant factors for the destination region. In most cases, measurable (quantitative) variables change over time, while qualitative factors, represented by dummies variables, are most often constant over time. It should be pointed out that, the inclusion of any dummy variables implies that the coefficient results must be interpreted in comparison with the excluded category of the corresponding variable. The terms μ_{ijt} represents the individual effect, specific to each pair of regions, which can be analysed as fixed over time or as random. Finally, we assume that the unexplained in the model part of the variation in migration flow is represented by the error term ε_{ijt}. Table 2.7. presents the definitions and data sources of all variables that were used in the econometric modelling.

Table 2.7: Explanatory variables considered in the econometric analysis of region to region migration flows

Variable	Definition	Source
Pop_o_{it}	population of the origin NUTS2 region in year t	Eurostat
Pop_d_{jt}	population of the destination NUTS2 region in year t	Eurostat
$lagGDPpc\_o_{it}$	lagged regional gross domestic product per capita of the origin NUTS2 region in year t – EUR	Eurostat
$lagGDPpc\_d_{jt}$	lagged regional gross domestic product per capita of the destination NUTS2 region in year t – EUR	Eurostat
$dist_{ij}$	the orthodromic distance between the centroids of the regions weighted by the population of the municipalities, corrected for enclaves, i.e. if the average distance of the points of a circle with an area equal to that of the larger of a given pair of regions from the centre of that circle (the average internal distance of the larger NUTS) is greater than the orthodromic distance between the centroids it replaces it.	IGSO PAS
$domestic_{ij}$	0- international; 1 – domestic	IGSO PAS
$lag\_disp\_inc\_o_{it}$	lagged disposable income of private households of the origin NUTS 2 region in year t	Eurostat
$lag\_disp\_inc\_d_{jt}$	lagged disposable income of private households of the destination NUTS 2 region in year t	Eurostat
$pop\_den\_rel_{ijt}$	the ratio of the population density of the NUTS 2 destination region to the population density of the NUTS 2 origin region in year t	Eurostat
$new\_eu\_o_{it}$	0 – the origin region in the “old” EU member country and EEA country; 1 – the origin region in new EU member country;	IGSO PAS
$non\_euro\_rel_{ijt}$	0 - both regions in Eurozone countries or in the same country; 1 - at least one region in a country that is not part of the Eurozone	IGSO PAS
$schen\_rel_{ijt}$	0 - at least one region is not in the Schengen area; 1 - both regions in the Schengen area or	IGSO PAS
$temp\_rel_{ij}$	difference in multi-year average temperature: average temperature in destination region minus average temperature in origin region	COPERNICUS
$precip\_rel_{ij}$	ratio of the multi-year average precipitation: average precipitation in destination region divided by average precipitation in origin region	COPERNICUS

Variable	Definition	Source
$island\_o_i$	0 - no "island" origin region, 1 – the island origin region	IGSO PAS
$island\_d_j$	0 - no "island" destination region, 1 – the island destination region	IGSO PAS
$outer\_o_i$	0 – not the outermost origin region, 1 – the outermost origin region	IGSO PAS
$outer\_d_j$	0 – not the outermost destination region, 1 – the outermost destination region	IGSO PAS
$language_{ij}$	0 - same language, 1 - same group but different language, 2 - different group. If more than one language in a region then the similarity of the most similar was taken into account.	IGSO PAS
$metro\_o_i$	0 - no MEGA in the origin region; 1 - MEGA 1 or 2 or 3 or 4 is in the origin region	IGSO PAS
$metro\_d_j$	0 - no MEGA in the destination region; 1 - MEGA 1 or 2 or 3 or 4 is in the destination region	IGSO PAS
$urban\_o_i$	2 - largest population of origin region lives in NUTS 3 "predominantly urban", 1 - largest population lives in NUTS 3 "intermediate", 0 - largest population lives in NUTS 3 "predominantly rural"	Eurostat, rural-urban typology for NUTS 3 based on DEGURBA indicator adjusted by IGSO PAS
$urban\_d_j$	2 - largest population of destination region lives in NUTS 3 "predominantly urban", 1 - largest population lives in NUTS 3 "intermediate", 0 - largest population lives in NUTS 3 "predominantly rural"	Eurostat, rural-urban typology for NUTS 3 based on DEGURBA indicator adjusted by IGSO PAS
$h\_edu\_o_{it}$	percentage of people with tertiary education in the 25-64 age group in the origin region	Eurostat
$rd\_exp\_o_{it}$	R&D expenditures the origin region	Eurostat
$agr\_sh\_o_{it}$	share of agriculture in the GVA of the origin region	Eurostat
$lab\_mar\_long_{ijt}$	long-term opening of the labour market for a given relationship and regions within the same country	IGSO PAS
$lab\_mar\_med_{ijt}$	medium-term opening of the labor market on a given relationship	IGSO PAS
$lab\_mar\_short_{ijt}$	short-term opening of the labor market on a given relationship	IGSO PAS
$unempl\_o_{it}$	percentage of unemployed in the origin region	Eurostat
$unempl\_d_{jt}$	percentage of unemployed in the destination region	Eurostat
$emp\_y\_o_{it}$	share of young people in total employment in the origin region	Eurostat
$emp\_sen\_o_{it}$	number of seniors per 100 employees of working age in the origin region	Eurostat
$emp\_w\_o_{it}$	percentage of female employees in the origin region	Eurostat
$emp\_mobil\_o_{it}$	employment of persons of mobile age in the origin region	Eurostat

Source: authors' elaboration.

As explained earlier, not for all variables defined above full sets of observations were available for all regions analyzed. Therefore, in some model specifications, the number of observations differed significantly from those where the explanatory variables were better represented.

The panel econometric model defined in this way, based on the gravity model, can be estimated by various methods (Khan and Hossain, 2010; van Bergeijk and Brakman, 2010) among which the simplest solution is OLS estimation of the linearized model by determining the natural logarithms of both sides of the regression. The choice of method depends on whether we face one of three important problems in the set of selected variables. The first potential problem arises from the fact that when we assume the presence of time-fixed individual effects for a given pair of regions there is no possibility to estimate the effects of time-constant factors, that is, many important explanatory variables in the model including the distance. The second problem is related to the potential endogeneity of explanatory factors. While the third difficulty arises from the fact that the zeros may appear in the dataset indicating that in some years there are no migration flows between certain regions. Then it is not possible to count the logarithm from the dependent variable, so such a model cannot be estimated in the classical way. A solution to the first and second problems was first proposed by Hausman and Taylor (1980) and developed by Baltagi and Khanti-Akom (1990) and Stock, et al. (2002). The so-called Hausman-Taylor estimator for panel data with random individual effects takes into account the division of explanatory variables into time-varying and constant factors. Moreover, in both sets one can identify factors that are known to be endogenous to the dependent variable. On the other hand, regarding the third problem, the so-called zero-flows problem, Santos and Tenreyro (2006) proposed the Pseudo Poisson Maximum Likelihood (PPML) estimator, which allows to estimate the gravity model without computing the logarithm of the dependent variable. Moreover, in this model it is possible to control the heterogeneity of the analyzed units. In the study presented here, where the dependent variable was migration flows between all regions of Europe in some cases there were zero flows. These cannot be considered random and treated as if they were missing observations, because the lack of migration between given regions may be due to the influence or lack of influence of specific explanatory factors. Therefore, the PPML estimator proposed by Santos and Tenreyro (2006) was used to estimate the analyzed relations. The models are then estimated in the following form:

$$M_{ijt} = \exp\left[\beta_0 + \beta_1 Pop_{it} + \beta_2 Pop_{jt} + \beta_3 lagGDPpc_{it} + \beta_4 lagGDPpc_{jt} + \right.$$
$$\left. + \beta_5 DIST_{ij} + \beta_6 domestic_{ij} + \beta_7 X_{it} + \beta_8 X_{jt} + \beta_9 D_i + \beta_{10} D_j + \mu_{ijt}\right] + \varepsilon_{ijt}\ .$$

It should also be noted that in order to obtain reliable results of the individual significance test of statistical parameters, robust standard errors were determined allowing to solve the problem of not meeting the assumptions about homoscedasticity and lack of autocorrelation in error term.

In order to verify if the identified explanatory factors are the same for all types of regions, in the next step of the analysis, econometric modeling of the model defined above was conducted for different groups of regions. The divisions were applied in which the following groups of relationships were included:

- **origins** are border regions and non-border regions;
- **destinations** are border regions and non-border regions;
- **origins** are regions in the new EU member states;
- **destinations** are regions in the new EU member states;

- **origins** are regions in the “old” EU member states;
- **destinations** are regions in the “old” EU member states;
- **origins** are regions with metropolitan areas (METRO 1 or 2 or 3 or 4) and regions without metropolitan areas;
- **destinations** are regions with metropolitan areas (METRO 1 or 2 or 3 or 4) and regions without metropolitan areas;

and separately for:

- migration between regions **within countries** (domestic);
- **international** migration between European regions.

All estimation results for each group are presented in separate tables.

3 Results

3.1 Country-to-country flows

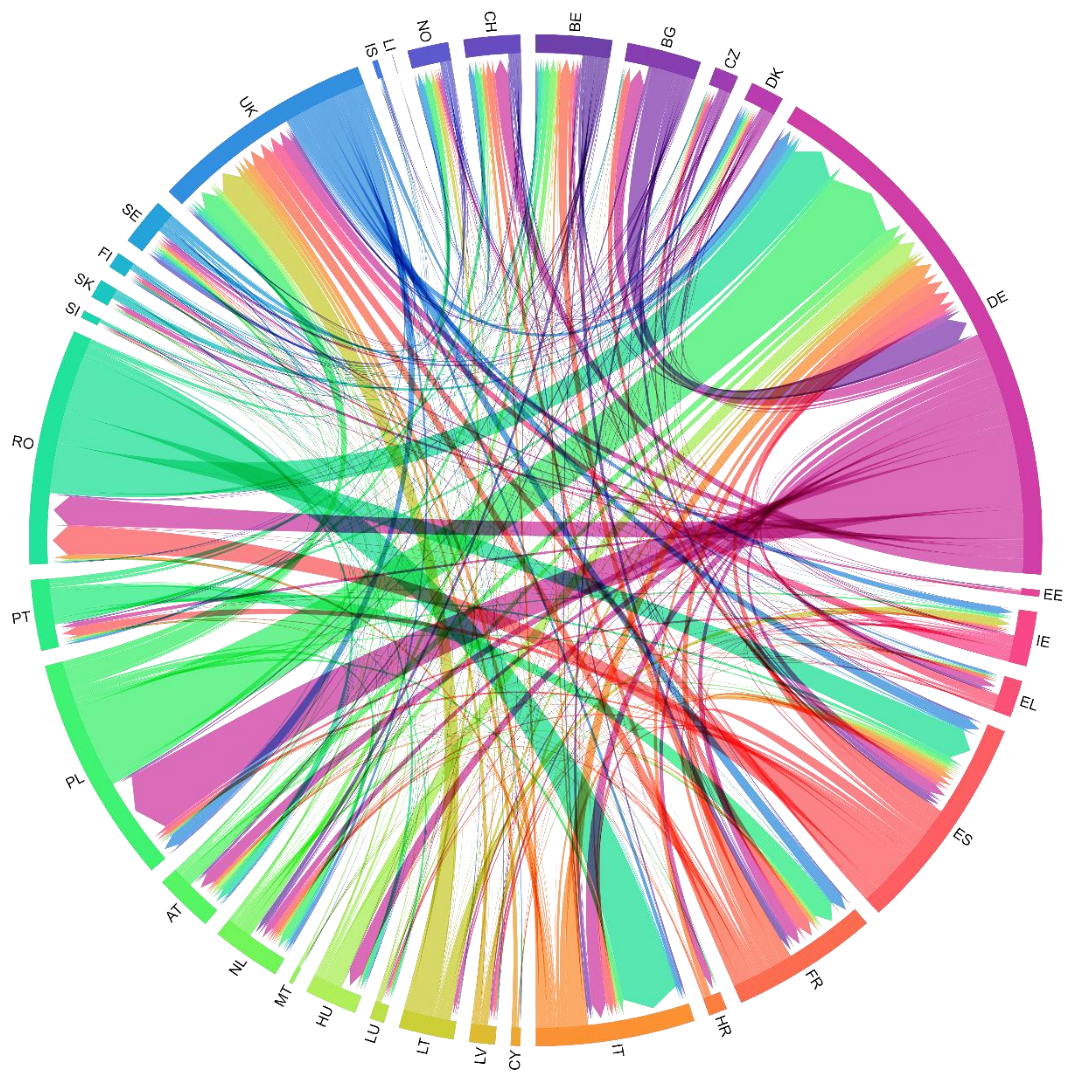
3.1.1 Description of results

The results of analyses of migration flows at the country level are presented in terms of: (1) migration inflow and outflow between ESPON space countries in the years 2010 and 2018; (2) dynamics of external flows between the periods 2010-2013 and 2015-2018, along with relation of domestic to external flows, and (3) migration inflow and outflow by selected countries.

For 2010 Germany clearly dominated for migration inflows and outflows alike (Fig. 3.1) (Germany accounted for more than 20% of all the C2C flows within the ESPON space). In terms of both migration inflows and outflows (2010-2018) the fundamental direction of flows was to Poland, followed by Romania and Bulgaria (Table 3.1). Coming in second place was Romania (8.9% of all flows), with a strong prevalence of outflows over inflows. Outflow dominated inflow also with Poland (8.6%), and Germany clearly dominated in the structure of migration flows. In 2010 large shares of flows characterized also the UK (8.5%), with domination of inflows, and Spain (8.3%), with a balance between inflows and outflows. The smallest magnitudes of flows in 2010 corresponded, of course, to the countries with the smallest populations (Liechtenstein – 0.0%, Malta – 0.2%, Iceland – 0.2%, Slovenia – 0.3%, Estonia – 0.3%, Cyprus – 0.4%).

In 2010 the highest intensity of migrations, in absolute values, among ESPON space countries was observed for the flows from Poland to Germany (and vice versa), from Romania to Italy and Germany, from Germany and Spain to Romania, and from Bulgaria to Germany. Also notable were countries that, despite their small population potential, featured relatively large migration flows. Chief among these were Luxembourg (481 persons per 10,000 inhabitants – accounting for the sum total of flows), Iceland (291 persons per 10,000 inhabitants), and Lithuania (284 persons per 10,000 inhabitants). The lowest intensity of flows in relation to population was found in France (37 persons per 10,000 inhabitants), Czechia (37 persons per 10,000 inhabitants) and Italy (43 persons per 10,000 inhabitants). For the majority of countries the value of this indicator was above the mean for the ESPON space (83 persons per 10,000 inhabitants).

Figure 3.1: Migration flows, 2010

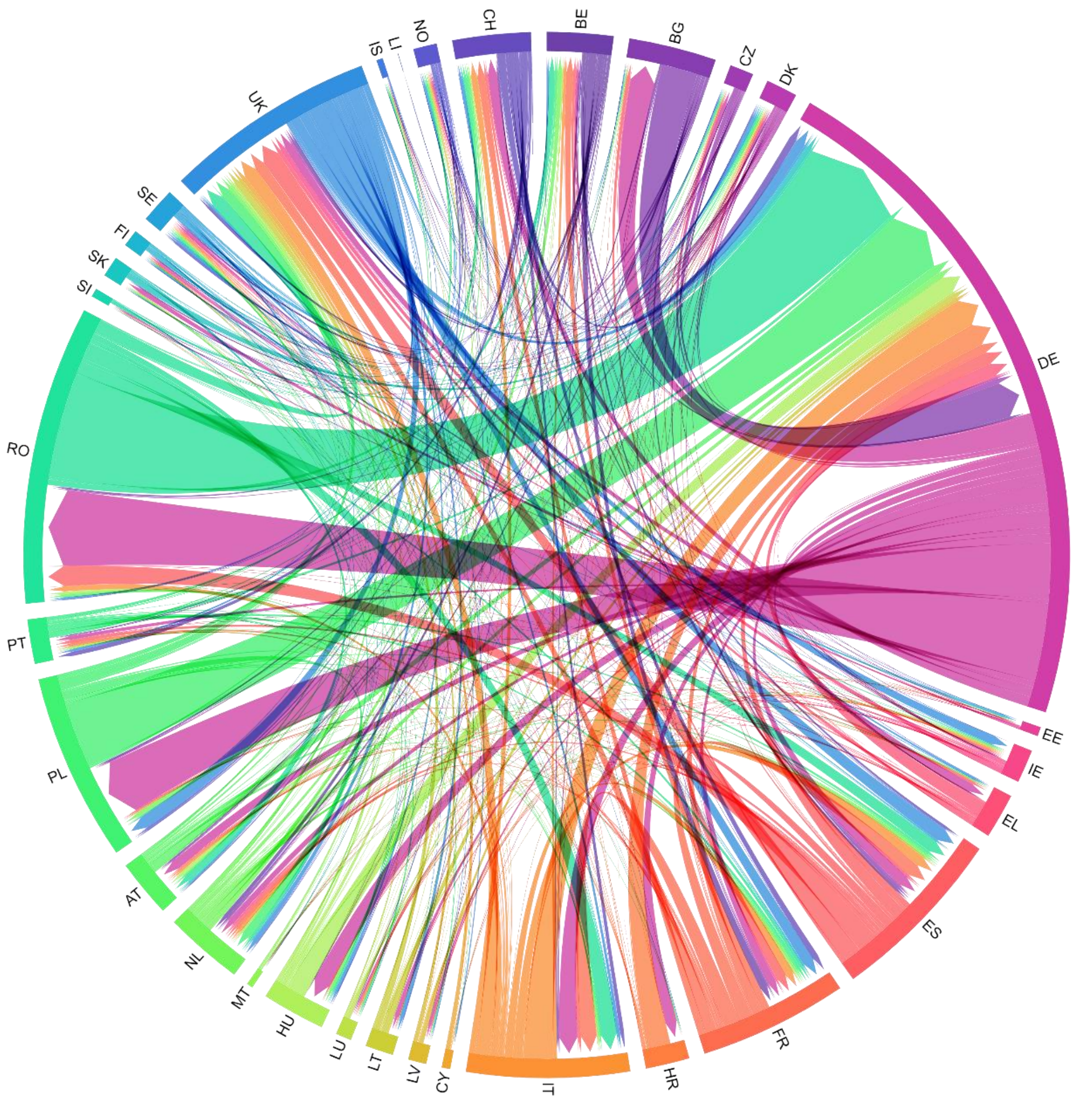


In the last year analysed (2018) the same countries retained the majority of flows inside the ESPON space, but their respective shares changed. Germany accounted for more than 25% of all flows (with inflows exceeding outflows). There was an increase in Romania's share of migration flows (10.9%), but in this case the difference between outflow and inflow decreased. The share of the UK in the total magnitude of flows in 2018 was 7.8%, with inflow clearly bigger than outflow (although this difference slightly diminished). The subsequent slots in the ranking were taken up by Poland (6.9%) and Spain (6.5%), both with a slight domination of outflows over inflows (in the case of Spain the respective proportions remained unchanged, while for Poland the difference decreased). The lowest magnitudes of flows were noted for the same countries as in 2010, with a slight drop of the share of Cyprus.

By the year 2018 the main directions of the biggest migration flows had not changed much. The biggest were the flows from Romania and Poland to Germany, and in the reverse direction, i.e. from Germany to Romania and Poland. As in 2010, the lowest values of the migration flow indicator were to be found in France (49 persons per 10,000 inhabitants), Czechia (50 persons per 10,000 inhabitants), and Italy (59 persons per 10,000 inhabitants). The indicator's highest values were noted for Luxembourg (581 persons per 10,000 inhabitants), Iceland (413 persons per 10,000 inhabitants), and Romania (336 persons per 10,000 inhabitants). The average value of the flow indicator for the entire ESPON space increased to 114 persons per 10,000 inhabitants.

A decrease in the value of the flow indicator was noted for five countries: Liechtenstein, Ireland, Norway, Portugal, and Sweden. For the remaining countries we observed an increase in the intensity of flows. The biggest was in Croatia, followed by Romania, Iceland, Malta, Bulgaria, and Luxembourg.

Figure 3.2: Migration flows, 2010



In the years 2010-2018 the highest concentration of migration outflows and inflows alike was observed for Croatia (Table 3.1). The vast majority of these migration flows (78.9% of outflows and 78.8% of inflows) was directed towards and from Germany. Croatia is followed in concentration of migration inflow by Poland, here Germany again dominating (71.6% of migration inflow comes from Germany, largely because of return and pendulum migrations). In 2018 Poland was among the countries with the highest share of inflows from and outflows to Germany (Fig. 3.3). High inflow concentration was also observed for Bulgaria, with, again, the key role of Germany and, to a much lesser extent, Spain. The lowest concentrations relative to migration inflow — close to null — were in the countries with an open migration policy: i.e. Sweden, Denmark, the UK, and the Netherlands.

As for concentration of migration outflow, the top spot was occupied by the aforementioned Croatia. The next ranks were occupied by Cyprus, with domination of the UK (73.6%), although, given their absolute values, these flows were hardly visible in the totality of inflows to the UK (Fig. 3.4), then Bulgaria (migration outflow to Germany at 67.0%) and Poland (also with domination of outflow to Germany, accounting for 65.3% of the respective total). The lowest values of the migration outflow concentration indicator were noted for Denmark, the UK, and Sweden.

In the dynamic perspective the changes were analysed from the beginning to the end of the period considered, to avoid randomness in the assessment of changes. Concerning absolute migration inflows, the biggest increases were observed for Germany, Romania, and the UK. Decreasing migration inflow, in both absolute and relative terms, was observed for four countries (Liechtenstein, Italy, Belgium, and Norway), with the biggest decrease noted for Norway. In percentages, however, the biggest upward changes in the migration inflows were observed for the small countries, i.e. Estonia (349.6%), Malta (209.6%), and Iceland (189.8%).

As for changes in migration outflows, the biggest ones in absolute terms were observed also for Germany and Romania, followed by Italy, France, and the UK. In percentage, the biggest increases occurred for Croatia (325.2%), Italy (165.5%), Switzerland (162.2%), and Finland (160.6%).

In general, a positive net migration balance (the difference in the sums of inflows and outflows in the years 2010-2018) was noted for the countries of the “old EU” (including Switzerland, Norway, and Iceland), and a negative one for the countries of the “new EU” (but also for four countries of the “old EU”, namely Italy, Greece, Portugal, and Spain). The biggest positive net balances were observed for Germany, the UK, and Switzerland, while the biggest negative balances were observed for Romania, Poland, and Spain.

Eight countries presented bigger external than domestic migrations in the years 2010-2018: namely, Lithuania, Croatia, Bulgaria, Poland, Portugal, Slovakia, Ireland, and Romania (Table 3.1). On the other hand, the biggest ratios of domestic to external migrations were noted for the UK (9.06), Finland (5.77), and Sweden (4.65).

The magnitudes of migration-based associations for the selected countries are presented as inflow and outflow, expressed as a percentage of the flows between the countries involved (2018). For Germany the dominating direction of connections (both ways) was towards Poland

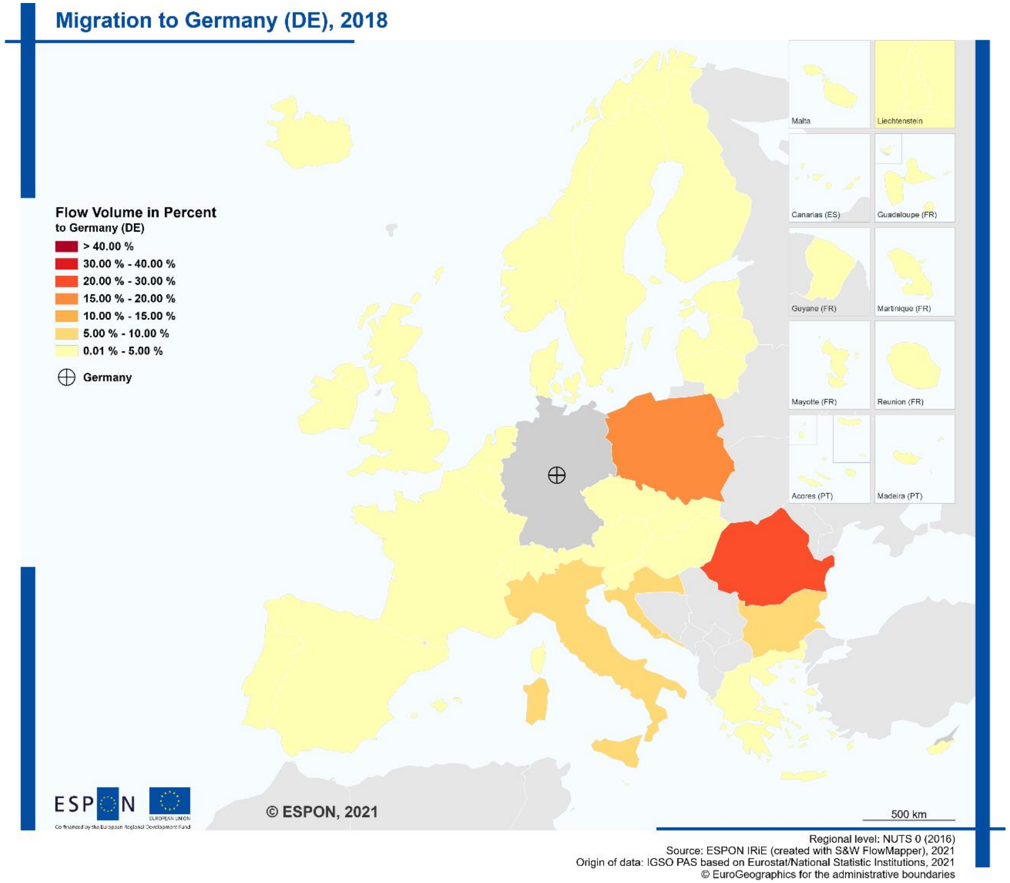
and Romania, followed by Bulgaria and Italy (Fig. 3.3). For the UK the dominant country in flows (both ways) was Spain. For inflows Italy and Romania were notable, while for outflows, alongside Spain, the main countries were France, Germany, and Poland (Fig. 3.4). Finally, the migration inflow to Italy was dominated by Germany, France, Romania, and Switzerland. The primary direction of migration outflows from Italy was towards Germany, followed by the UK and Spain (Fig. 3.5).

Table 3.1: Basic characteristics of C-to-C flows

Country	Inflow		Outflow		Balance	Relation of internal to external migrations	Top 3 (2010-2018)						Concentration index HHI (2010-2018)	
	Difference	Change	Difference	Change	Sum		Inflow countries (% of total inflow)			Outflow countries (% of total outflow)				
	2015-2018 / 2010-2013	2015-2018 / 2010-2013	2015-2018 / 2010-2013	2015-2018 / 2010-2013	2010-2018	2010-2018	1	2	3	1	2	3	Inflow	Outflow
AT	31272	111,5	11368	106,0	207896	1,82	DE (28.1)	RO (15.8)	HU (13.0)	DE (38.3)	RO (10.5)	HU (10.2)	1368	1856
BE	-6547	97,9	34495	115,5	170420	2,14	FR (20.7)	NL (13.5)	RO (12.4)	FR (27.1)	NL (14.0)	DE (9.4)	1048	1235
BG	82637	140,9	117402	134,5	-360216	0,31	DE (68.0)	ES (15.5)	AT (2.3)	DE (67.0)	ES (6.1)	UK (4.2)	4894	4593
CH	16441	104,5	111465	162,2	330276	1,23	DE (24.2)	FR (16.5)	IT (14.7)	DE (25.6)	FR (17.1)	IT (11.0)	1286	1301
CY	4159	127,3	12116	123,5	-90821	0,00	UK (50.5)	DE (10.7)	FR (6.7)	UK (73.6)	ES (8.5)	DE (6.0)	2800	5538
CZ	21750	127,3	18176	120,2	-17589	3,29	DE (33.6)	ES (18.2)	UK (8.5)	DE (44.8)	ES (15.5)	SK (5.4)	1663	2363
DE	884677	133,8	735096	140,7	2071276	2,71	PL (21.9)	RO (21.3)	BG (8.7)	PL (22.3)	RO (20.1)	BG (7.6)	1192	1134
DK	24078	118,6	13623	113,2	73043	3,35	SE (12.3)	PL (10.0)	RO (9.9)	SE (13.6)	DE (12.8)	NO (10.6)	707	747
EE	32152	349,6	2294	105,7	-29378	0,00	FI (47.9)	DE (11.8)	UK (6.3)	FI (53.1)	DE (11.5)	UK (6.3)	2555	3073
EL	24616	121,1	57173	129,6	-206620	1,07	DE (48.3)	ES (12.7)	UK (11.5)	DE (53.5)	UK (11.5)	ES (7.2)	2706	3120
ES	114295	119,4	-17217	98,0	-515744	1,94	RO (20.6)	UK (13.7)	DE (13.1)	RO (24.1)	DE (15.1)	UK (15.1)	1081	1249
FI	3591	105,1	29662	160,6	22103	5,77	EE (29.7)	SE (15.5)	DE (12.2)	EE (21.0)	SE (19.4)	DE (16.0)	1390	1253
FR	42459	107,3	220562	148,7	59243	2,58	ES (15.2)	UK (15.1)	PT (13.3)	UK (17.7)	DE (15.0)	ES (12.6)	1052	1100
HR	52177	180,3	188599	325,2	-205330	0,35	DE (78.9)	AT (5.6)	SI (5.5)	DE (78.8)	AT (6.6)	SI (2.7)	6292	6269
HU	71140	139,7	29332	110,1	-209416	3,47	DE (62.1)	AT (9.3)	RO (7.6)	DE (62.5)	AT (12.2)	UK (6.7)	4049	4137
IE	37308	128,1	-21945	85,0	34361	0,85	UK (39.2)	LT (12.2)	RO (6.4)	UK (48.6)	DE (8.9)	FR (7.6)	1880	2598
IS	15920	189,8	-2283	87,2	19016	0,00	PL (28.0)	DK (16.1)	NO (9.0)	NO (20.0)	DK (19.3)	SE (17.0)	1305	1371
IT	-4225	99,3	307423	165,5	-71868	2,15	RO (39.4)	DE (21.8)	FR (9.9)	DE (35.5)	UK (12.7)	ES (9.9)	2217	1751
LI	-4	99,8	103	106,0	1383	0,00	CH (60.6)	AT (16.8)	DE (9.8)	CH (63.7)	AT (13.4)	DE (8.8)	4068	4330

Country	Inflow		Outflow		Balance	Relation of internal to external migrations	Top 3 (2010-2018)						Concentration index HHI (2010-2018)	
	Difference	Change	Difference	Change	Sum		Inflow countries (% of total inflow)			Outflow countries (% of total outflow)				
	2015-2018 / 2010-2013	2015-2018 / 2010-2013	2015-2018 / 2010-2013	2015-2018 / 2010-2013	2010-2018	2010-2018	1	2	3	1	2	3	Inflow	Outflow
LT	12846	117,2	-47934	79,0	-259777	0,28	UK (33.0)	DE (27.9)	IE (8.4)	UK (44.2)	DE (19.0)	NO (10.0)	2071	2533
LU	8419	112,2	19393	145,6	49807	0,00	PT (21.7)	FR (19.9)	DE (14.1)	DE (27.2)	FR (20.2)	PT (15.2)	1324	1605
LV	6486	110,5	-20065	82,6	-87361	0,00	DE (31.2)	ES (24.9)	UK (10.7)	UK (32.2)	DE (30.1)	IE (7.3)	1832	2087
MT	8980	209,6	1974	108,2	-25676	0,00	IT (20.3)	UK (10.3)	DE (9.5)	UK (52.2)	ES (18.6)	DE (5.0)	872	3138
NL	82464	129,0	63767	129,7	172660	3,86	PL (19.5)	DE (14.8)	BE (10.4)	DE (22.3)	BE (17.2)	UK (13.2)	958	1168
NO	-65314	65,2	26436	134,6	150128	3,05	PL (21.6)	SE (15.5)	LT (12.7)	SE (26.3)	DK (12.3)	PL (12.2)	1047	1215
PL	127290	120,8	-49548	95,3	-811788	0,41	DE (71.6)	UK (11.1)	NL (3.1)	DE (65.3)	UK (12.0)	NL (6.1)	5279	4470
PT	48955	133,1	-125943	63,0	-229393	0,57	ES (21.9)	DE (17.3)	FR (16.4)	FR (28.9)	DE (14.9)	CH (13.5)	1326	1503
RO	366067	161,4	494400	145,1	-1283897	0,98	DE (56.1)	ES (26.7)	IT (5.4)	DE (49.1)	IT (17.5)	ES (9.7)	3908	2916
SE	11384	106,4	-11792	91,5	117835	4,65	NL (12.8)	NO (10.0)	DE (9.4)	NO (18.5)	DK (13.3)	DE (12.5)	674	980
SI	3310	111,3	14341	137,2	-35002	2,34	DE (39.0)	HR (15.9)	AT (9.0)	DE (42.8)	AT (15.9)	HR (10.7)	2019	2289
SK	15901	123,9	9626	110,5	-50314	0,80	DE (50.0)	AT (12.9)	CZ (7.1)	DE (52.1)	AT (16.5)	HU (6.3)	2797	3096
UK	336675	136,9	189258	134,2	1010744	2,25	ES (12.0)	PL (11.5)	RO (11.1)	FR (14.2)	ES (13.6)	DE (11.9)	749	856
Total	2411358	125,1	2411358	125,1	0	1,96	DE (20.1)	RO (12.4)	PL (9.6)	DE (28.6)	UK (10.0)	RO (7.2)	877	1165

Figure 3.3: Migration flows between Germany and ESPON space countries, 2018



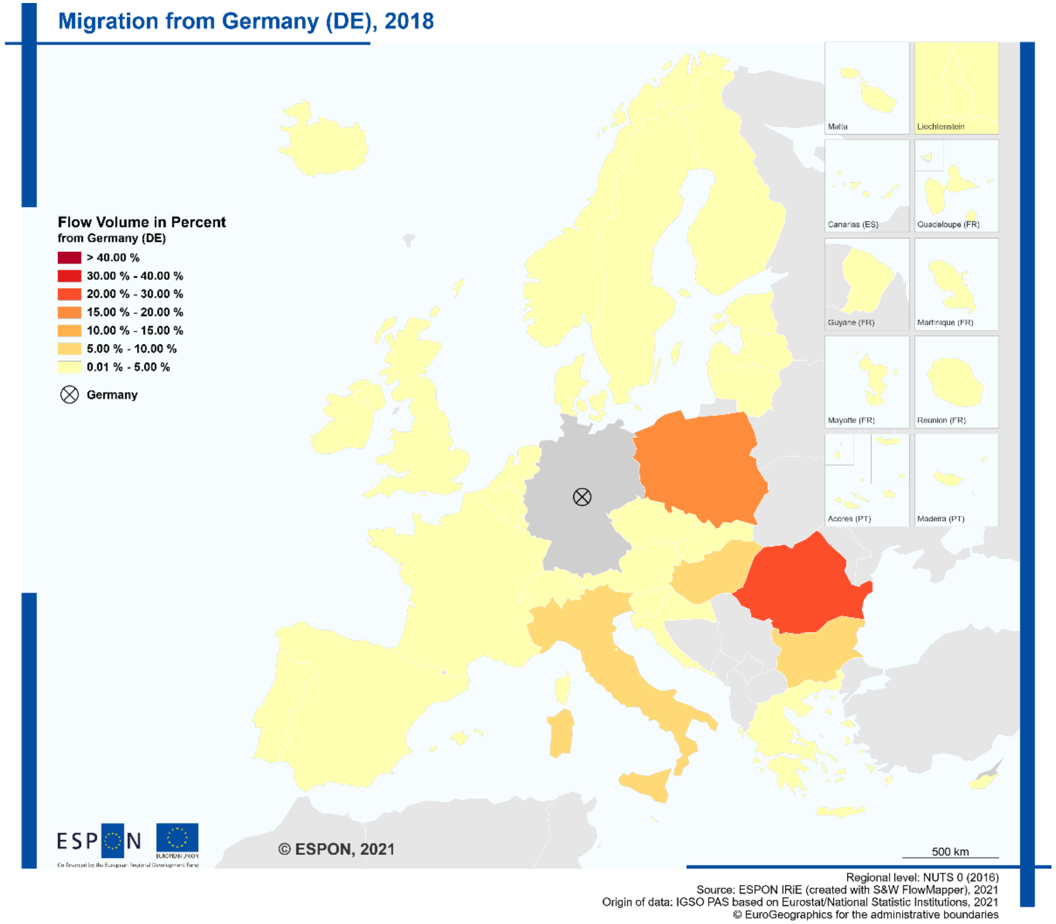


Figure 3.4: Migration flows between the UK and ESPON space countries, 2018

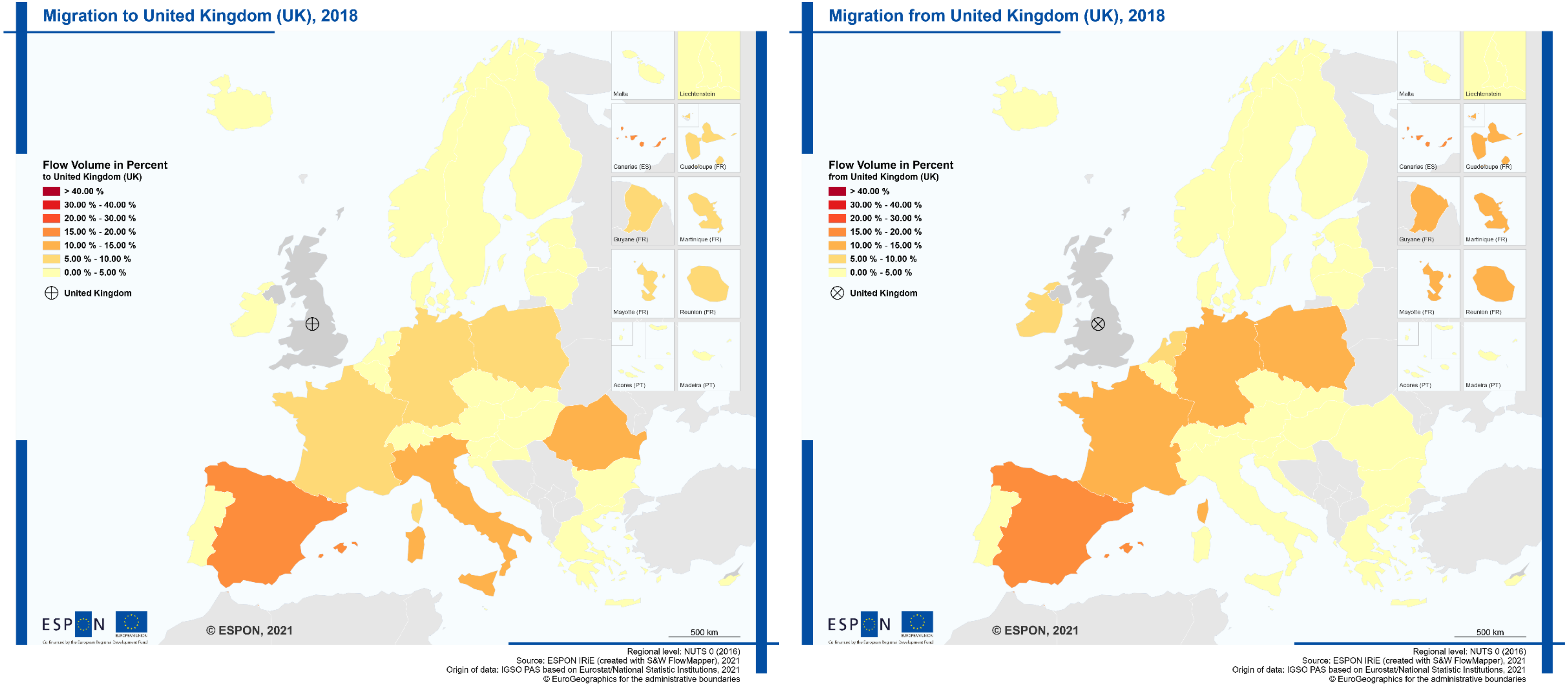
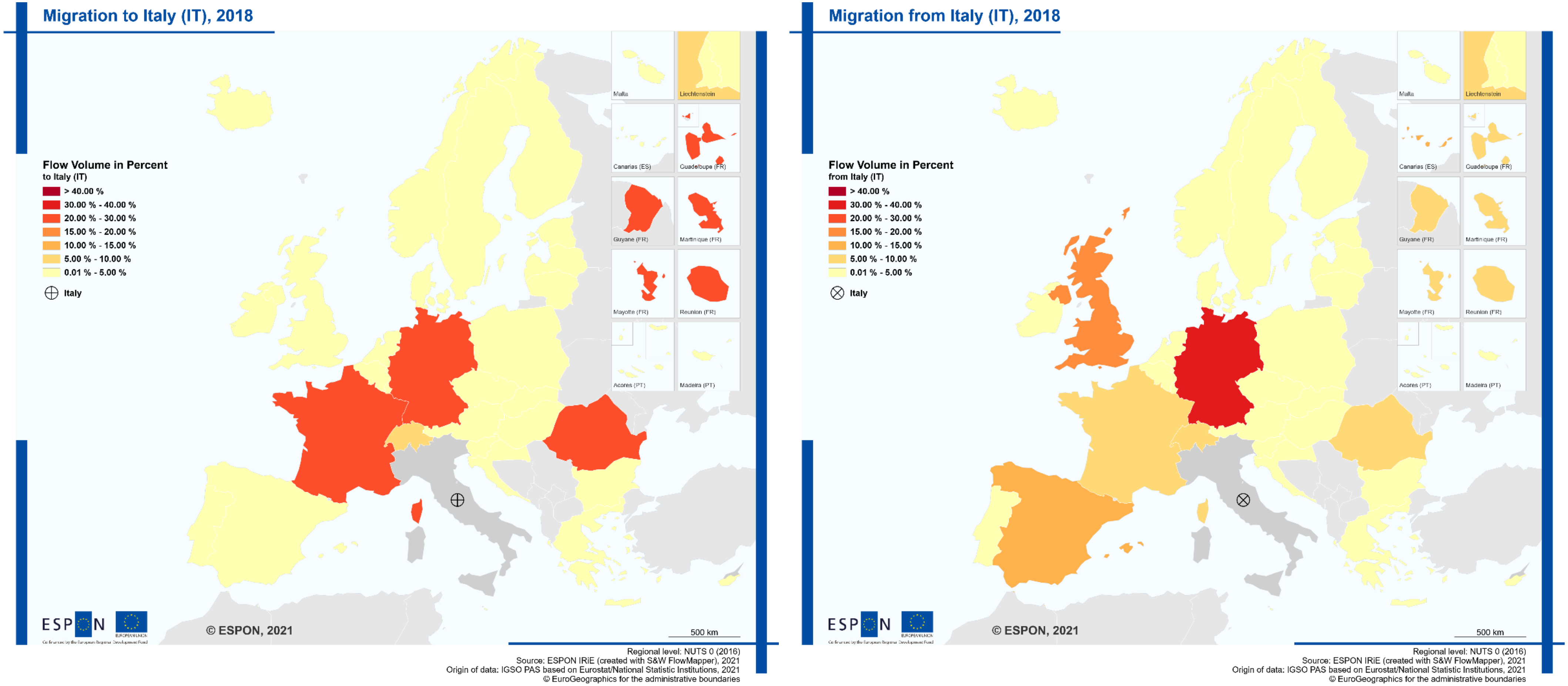


Figure 3.5: Migration flows between Italy and ESPON space countries, 2018



3.2 Region-to-region flows

3.2.1 Description of results

The project's most important objective is to identify regional flows. Regional migration flows are highly concentrated, as the 50% of regions with the fewest flows accounted for only 16% of total migration flows (R2R). On the other hand, the 10% of regions with the most flows accounted for as much as 40% (Fig. 3.6).

Figure 3.7 shows the 500 biggest and smallest interregional flows within the ESPON space countries for the year 2018. Migration flows at the regional level were characterised as follows:

- The biggest population flows were contained within a given country, so that only sporadically did interregional interactions cross state boundaries (e.g. between the Opolskie region in Poland and Düsseldorf in Germany, between some Romanian and German regions, and between Croatia — HR04, Kontinentalna Hrvatska — and certain German regions);
- The biggest interregional flows were generated by urban agglomerations along with their surroundings (Functional Urban Areas); and the biggest of these occurred with certain cities in Western Europe, chief among them (1) London, Birmingham, and Liverpool; (2) Amsterdam, The Hague, and Rotterdam; (3) cities in the Ruhr basin, the functional areas of Munich, Berlin, and Hamburg; (4) Vienna; (5) Copenhagen; (6) Helsinki; (7) Stockholm; and (8) Madrid (Fig. 3.7). It should be remembered, though, that such results depend also upon the magnitude of the NUTS 2 units in particular countries;
- Surprisingly big flows occurred between regions of Romania, not just in relation to Bucharest but also between other cities in Romania (e.g. in western Romania). We saw a similar situation in Finland, which featured a relatively low population potential. Such big flows might have resulted, on the one hand, from the inhabitants' high mobility or, on the other, from administrative conditioning (i.e. registering all kinds of movements, e.g. also for a short period of time);
- The biggest flows between the regions of Italy and of France were relatively limited in scale; similarly, interregional flows were virtually absent in Poland, Bulgaria, Portugal, and Slovakia (in all of which external migrations were decidedly higher than domestic regional migrations);
- Interregional flows between countries reflected primarily the "new EU" – "old EU" relation. Flows between regions in the countries of the "new EU" were virtually non-existent (Fig. 3.7). The lowest magnitudes of interregional flows (statistically below 1 person per year) occurred between the regions of Central Europe and the Balkans, and between the outermost regions (Ceuta, Melilla, Canary Islands, Azores, Aland Islands, etc.).

Figure 3.6: Concentration of migration flows, 2010-2018

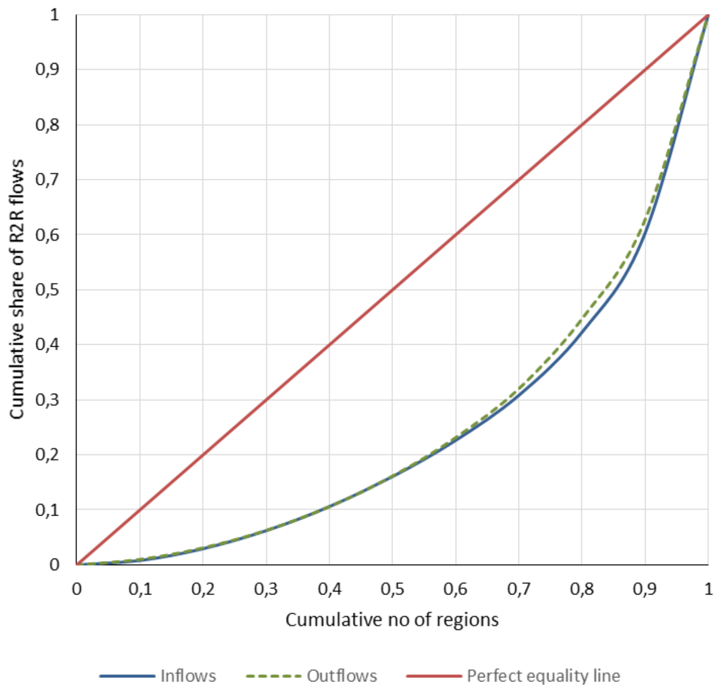
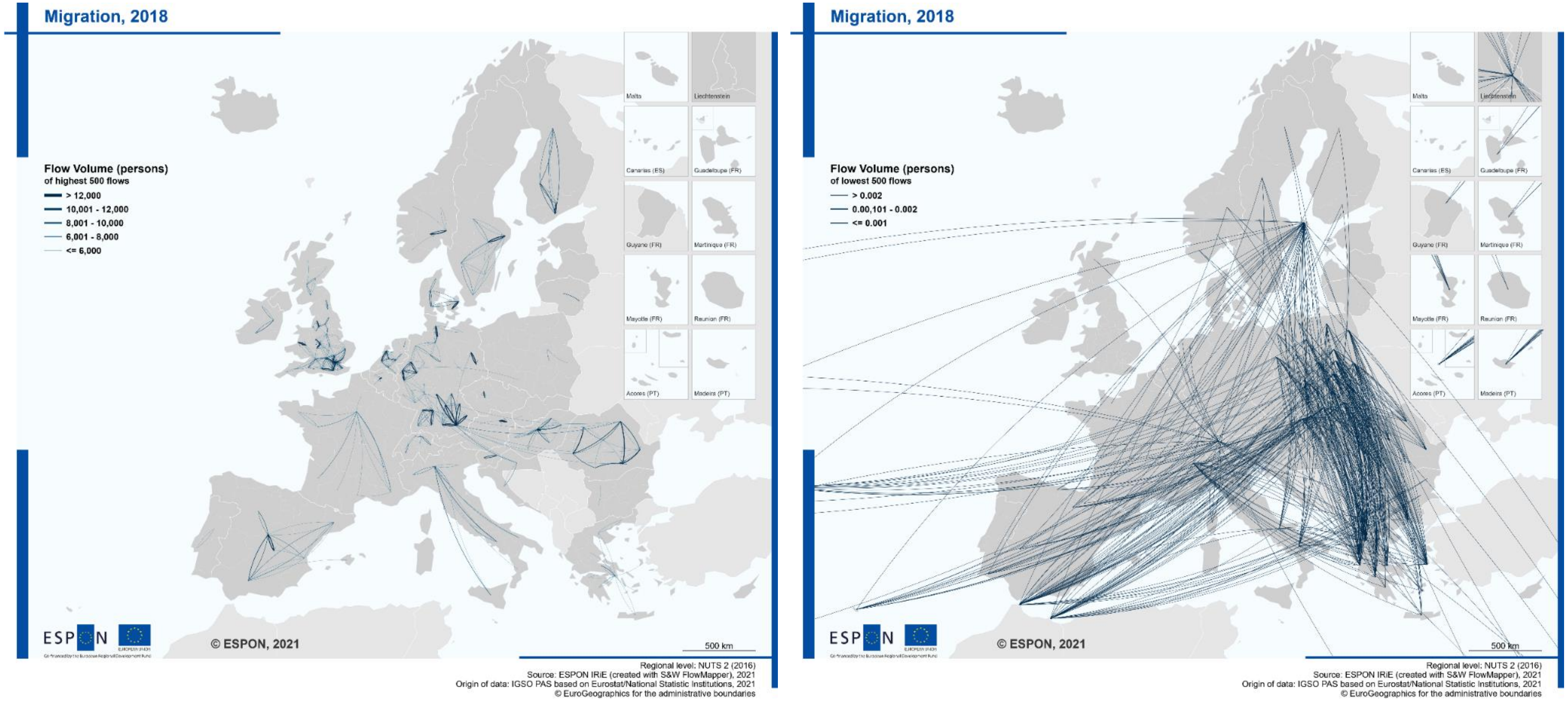
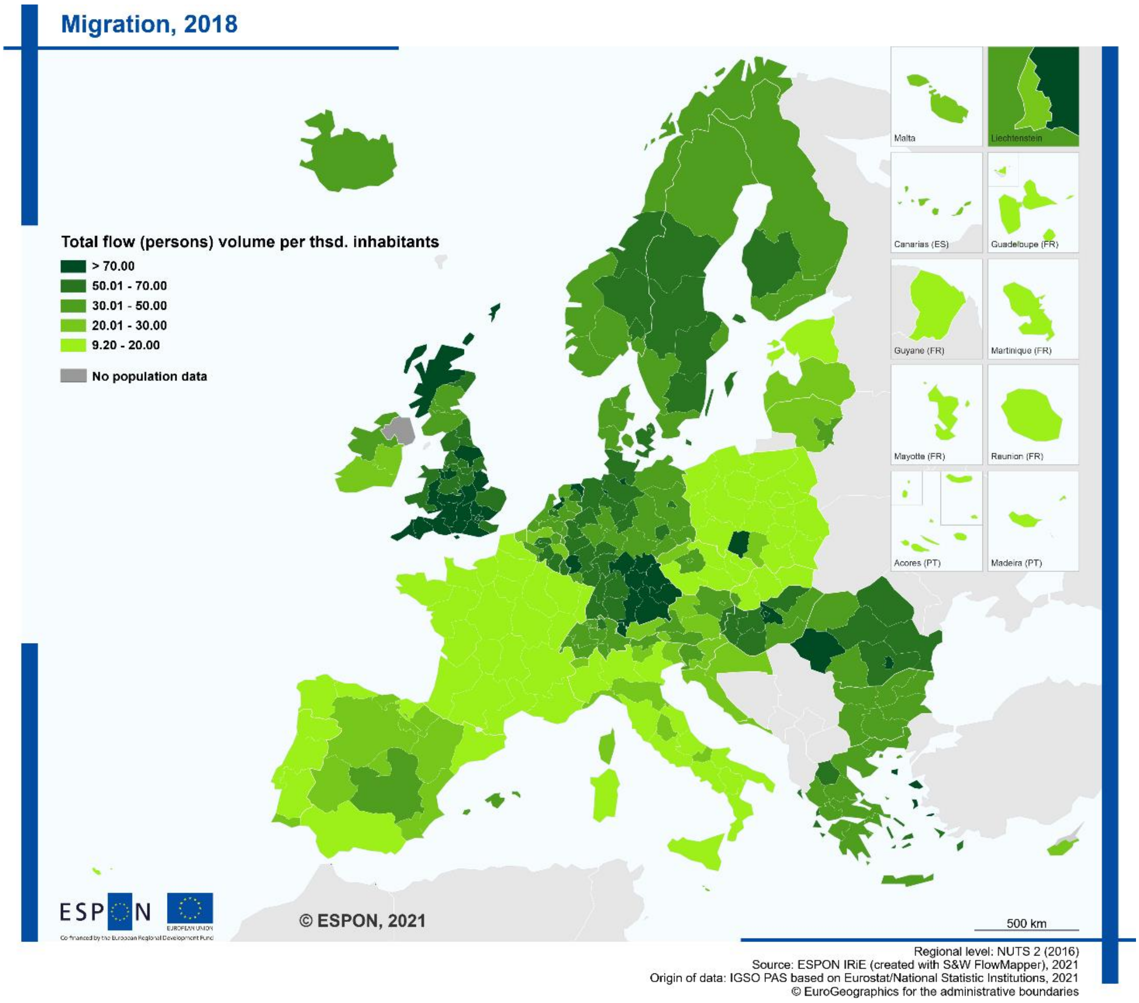


Figure 3.7: The smallest and the biggest 500 flows between regions of the ESPON space countries, 2018



The intensity of interregional migrations was highly differentiated within the ESPON space. In 2018 the highest magnitude of flows (inflows and outflows jointly) in relation to the population of a given region occurred in the UK, where in many regions the indicator's value exceeded 70 persons per 1,000 inhabitants (Fig. 3.8). A similar situation occurred in Bavaria, Germany. The Netherlands, Romania, Austria, Poland, and Hungary were in the single units. On the other hand, the indicator's lowest values (in 2018) were observed for all regions of France and for almost the entirety of Poland, Czechia, and Slovakia, as well as a large part of Italy. Regions of intensive migration inflow (the UK, Germany, Scandinavia) and of intensive migration outflow (Hungary, Romania, Bulgaria) were therefore clearly distinguishable.

Figure 3.8: Total volume of flows per 1,000 inhabitants, 2018



The years 2010-2018 saw a decrease in migration flows in the majority of regions of Spain and France, as well as in the southern part of Italy (Fig. 3.9). The biggest drops occurred in Spain (Castilla-La Mancha, Castilla y Leon), Italy (Lazio, Sicilia), and the UK (Northern Ireland). On the other hand, the greatest increase in migration inflow occurred in Germany, the UK, Belgium, the Netherlands, Hungary, and Romania. It ought to be remembered, though, that in these countries (and to the least extent in Romania) regional migration inflows were generated primarily by domestic migrations. (In these cases internal flows clearly exceed external flows.)

The greatest decreases in regional migration outflows were observed in the majority of the Spanish regions (e.g. Cataluña, Comunidad de Madrid), in the whole of Portugal, Ireland, Lithuania (the region with highest value in the study), and in southwestern Poland (Śląskie, Dolnośląskie) (Fig. 3.10). For international migrations this may be evidence of, on the one hand, a certain migratory saturation (decrease of migration potential of regions of origin) and, on the other, of a decrease in population mobility, resulting, in particular, from a decrease in wage differences between the respective countries and regions. The biggest increase in regional migration outflows was observed for the UK, Germany, Italy, Hungary, Croatia, Belgium, the Netherlands, and some of the Romanian regions, as well as individual metropolitan areas (e.g. Budapest, Copenhagen, Stockholm, Berlin, Athens). In these cases internal migrations and the accompanying processes (e.g. suburbanization) played an important role.

Figure 3.9: Change in incoming flows (external + internal), 2010-2018

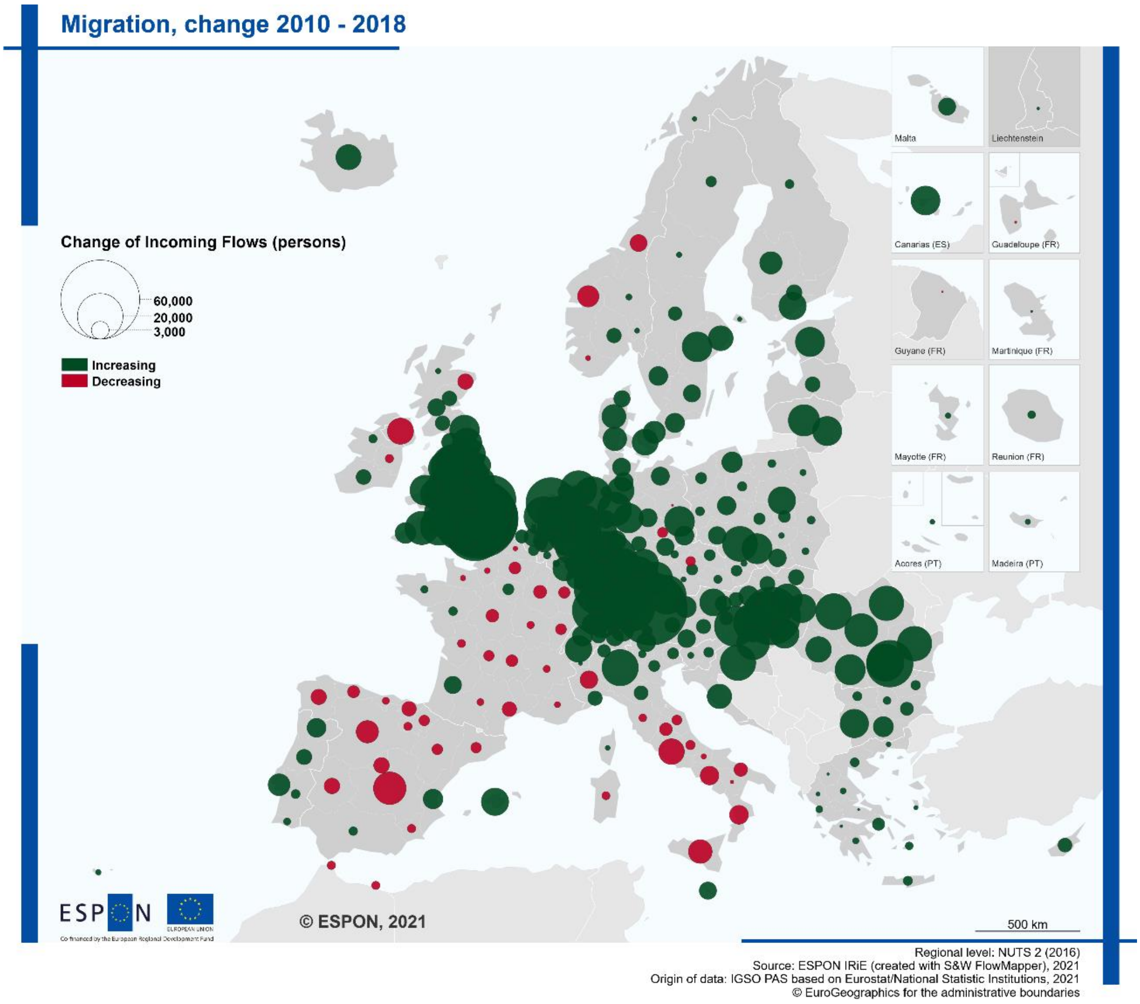
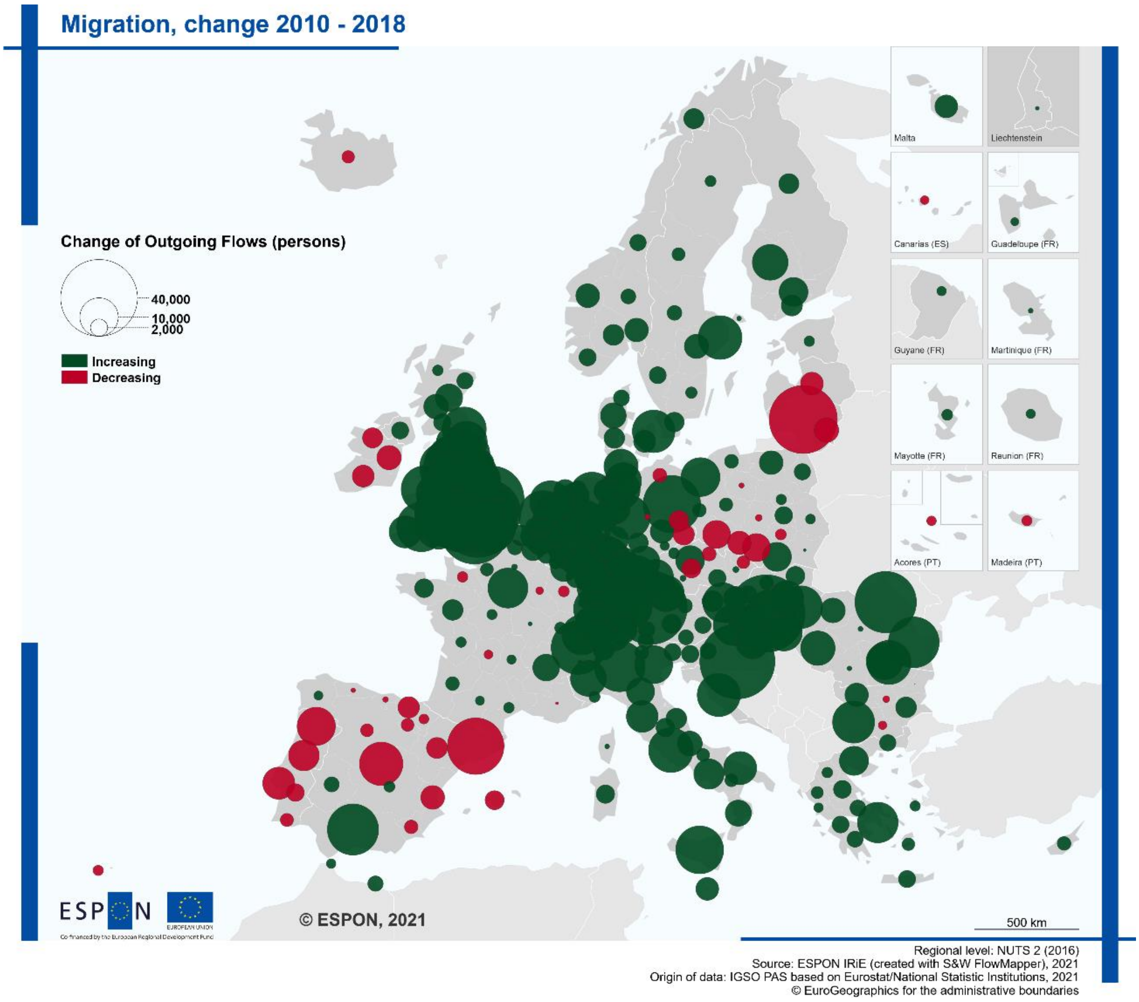


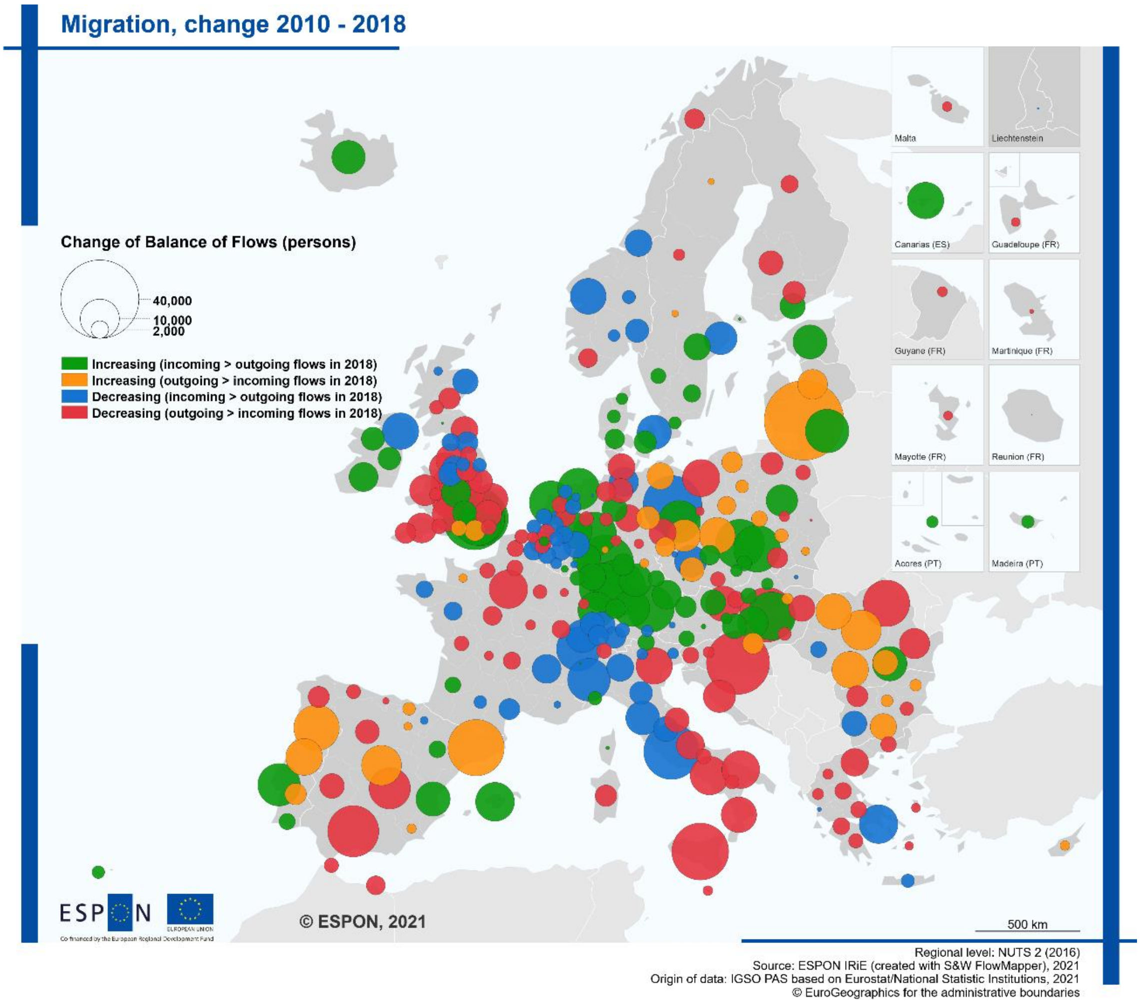
Figure 3.10: Change in outgoing flows (external + internal), 2010-2018



A positive net balance of flows over the years 2010-2018 with an upward tendency (increasing: incoming > outgoing flows) characterized primarily the majority of German regions (the biggest concentration of such regions) (Fig. 3.11). A similar tendency was observed for some urban agglomerations (e.g. Lisbon, Valencia, Warsaw, London, Bucharest, Helsinki). Yet with urban agglomerations we are dealing also with groups of cities, which featured a drop in net migration balance, which remained positive in 2018 (decreasing: incoming > outgoing flows). This applied in particular to Stockholm, Oslo, Athens, Rome, Copenhagen, and Berlin.

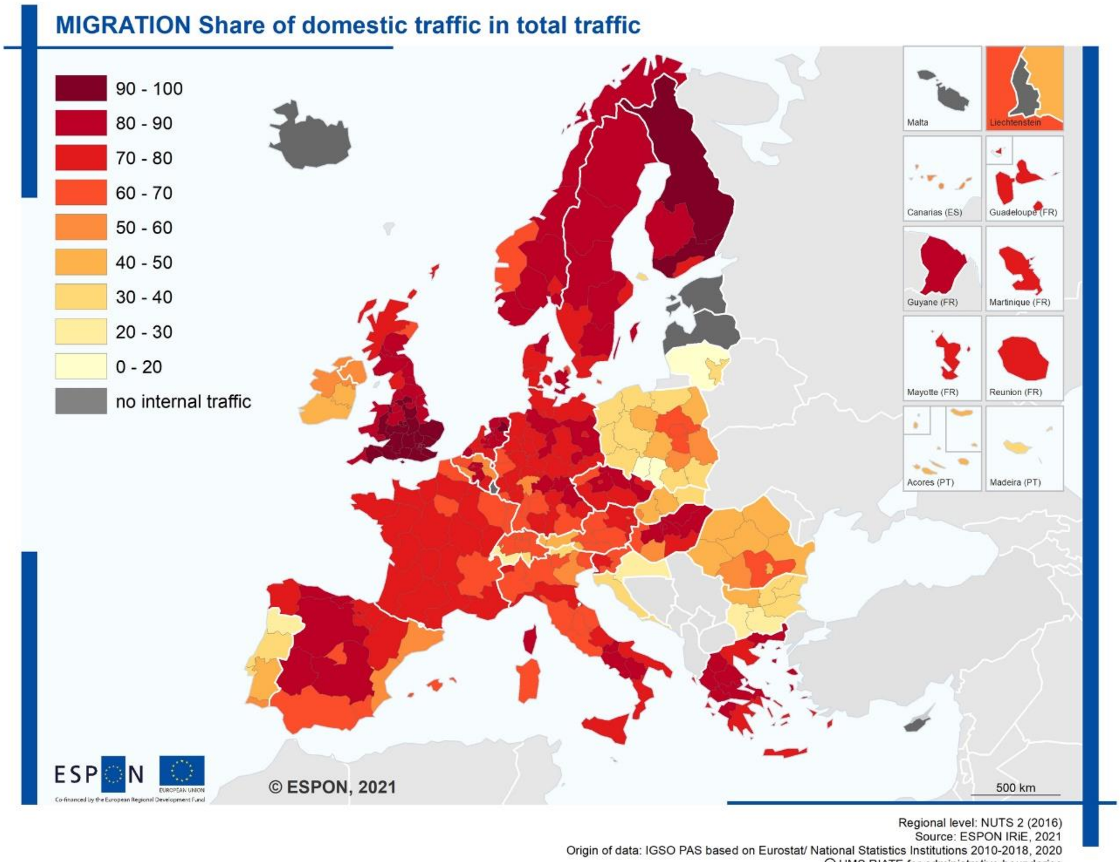
Many of the ESPON space regions presented a deepening negative net migration balance (decreasing: outgoing > incoming flows) (Fig. 3.11). This concerned, first of all, Croatia (Kontinentalna Hrvatska), southern Italy (especially Sicily), Andalusia in Spain, Budapest, Ile-de-France, and a significant number of regions in the UK. These regions feature a dynamic migration outflow. Still, the outflow's nature was quite diverse. Romania (Nord-Est region) presented job-related foreign migrations primarily. These could reflect suburbanization in Budapest, but in the regions of southern Italy they could reflect migration outflow towards the country's richer northern regions. The subsequent case is decreasing negative net migration balance (i.e. actual increase in the value of net migration balance), which was observed in, for example, in regions of Portugal, Spain (e.g. Catalonia), Romania, Poland, and Bulgaria. The most striking instance was Lithuania (LT02, Vidurio ir vakaru Lietuvos regionas), where the negative net migration balance decreased by more than 41,000 persons (from -60,200 in 2010 to -18,800 in 2018).

Figure 3.11: Change in balance of flows (external + internal), 2010-2018



One of the key influences on the indicators was the share of internal migrations in total bilateral regional migration flows (Fig. 3.12). Its significance resulted from the different natures (in conditioning, magnitude, structure) of the two kinds of movements. Moreover, the role of internal flows remained conditioned by the magnitude of the NUTS 2 units in the particular countries. Smaller units brought more local migration movements into the study. One must be aware of these limitations to interpret the results correctly. In formal terms, shares of internal migrations ranged from roughly 20% in some regions of Bulgaria, Portugal, and Poland to more than 90% in the Scandinavian countries and the UK.

Figure 3.12: Share of domestic migrations in total regional migration flows



Regional level: NUTS 2 (2016)
Source: ESPON IRiE, 2021
Origin of data: IGSO PAS based on Eurostat/ National Statistics Institutions 2010-2018, 2020
© UMS RIATE for administrative boundaries

Low shares were observed in certain countries of Central-Eastern Europe (Poland, Romania, Slovakia, Croatia, Bulgaria) and in Portugal. The values of this indicator were usually higher in the surroundings of large metropolises, but not in their city cores. This indicates the role of certain large centres, which, owing to their attractiveness, constituted for neighbouring areas the essential competitors to foreign migration destinations. At the same time cities were already participants in movements at the European level. Consequently, the share of domestic migrants was smaller for them.

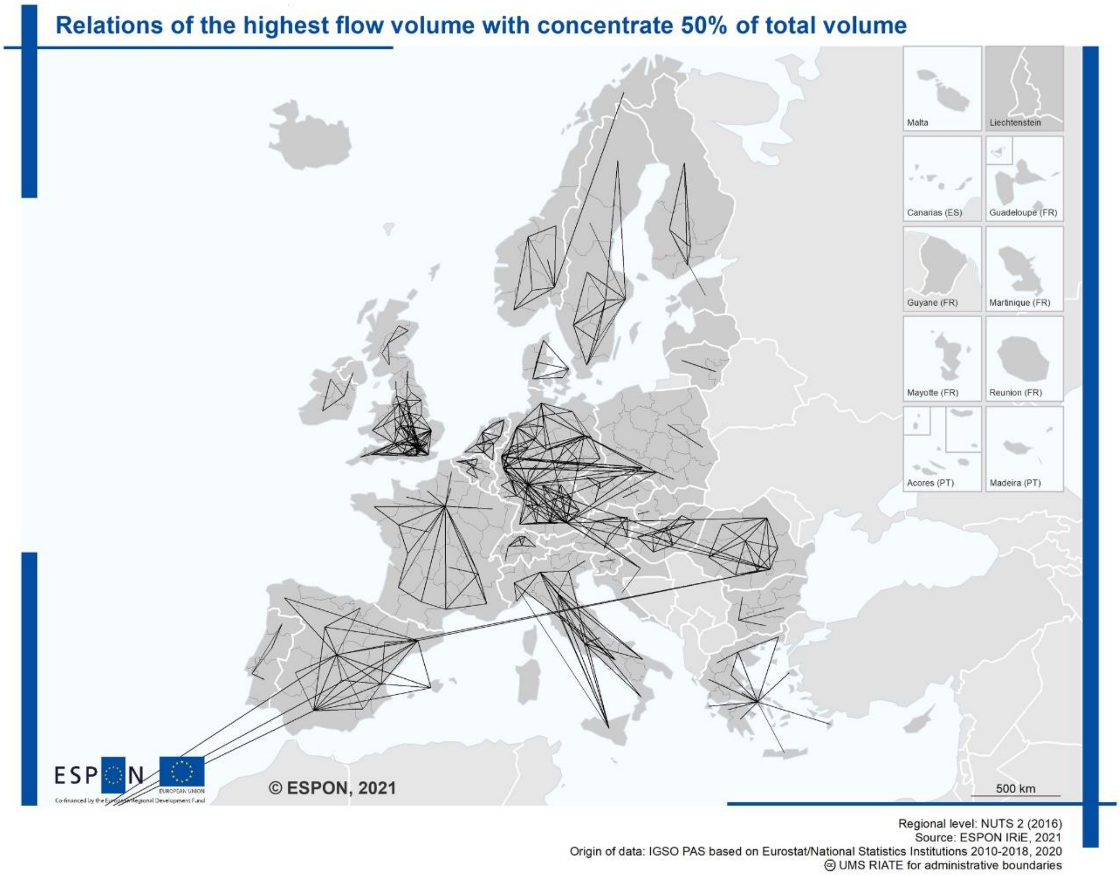
There were characteristic differences in the shares of domestic migrations within certain countries. These were visible in Poland (higher share of internal flows in the eastern part of the country), Spain (lower significance of such flows in the South: in regions attracting foreign migrants), Germany (more international migration flows in western lands), and also Italy (domestic migrations more important in the southern part of the country). It can be supposed that these differences resulted from such factors as: a) internal differentiation in development and human capital; b) participation in the inter-metropolitan exchange of highly skilled personnel; c) high residential attractiveness of some regions.

3.2.1.1 Distribution of the biggest migration flows

Most of the strongest migration-related connections in Europe (Fig. 3.13) were domestic connections, within countries. Their spatial pattern was — in certain countries — a consequence of the magnitude of the regions considered. Small units (especially in Germany and the UK) cause that the analysis at NUTS 2 level encompasses the major part of the total internal migration flows. These flows often include movements within metropolitan areas (whenever these were divided up by unit boundaries, as in the Ruhr Basin or Greater London). And so this may be an expression of, say, suburbanization. In other countries, such as France, Spain, Italy, and Poland, such migrations were not visible in the analysis. Despite these reservations, we can perhaps ascribe the dominance of domestic migrations to the persistence of the “border effect” in social relations, even between countries where integration is strong and long lasting. An instance was provided by the strongly ”internally closed” pattern of the biggest migration flows in Belgium and the Netherlands (despite more than 60 years that have passed since the Benelux community was founded). In this sense the study confirms, indirectly, the proposition of the persistence of boundaries, even when they are formally fully permeable (see Reietvald, 2001).

The most spectacular exceptions to the dominance of internal migrations were found in the relations between Romanian units and Germany and, in a couple of cases, with Spain, as well as in the relations of two regions in southern Poland with the western lands of Germany. In the latter case, though, the high intensity of international movements primarily concerned areas where the inhabitants had for many years enjoyed dual citizenship (the area of Opole, with German minority). Intensive migrations had been taking place there for many years before Poland’s accession to the EU. The consequence nowadays is return migrations (after the termination of professional careers), which resulted in a generally higher magnitude of flows. This was not observed in the poorest Polish, Romanian, and Bulgarian regions, where bigger migrations had appeared only after 2004 and 2007. Besides, the system of biggest flows included also relations between Croatia and Germany (Bavaria), Vienna and Munich, and Estonia and Finland (relatively significant, in light of the low population potentials of these two countries).

Figure 3.13: Biggest migration flows (50% of total migration movements)

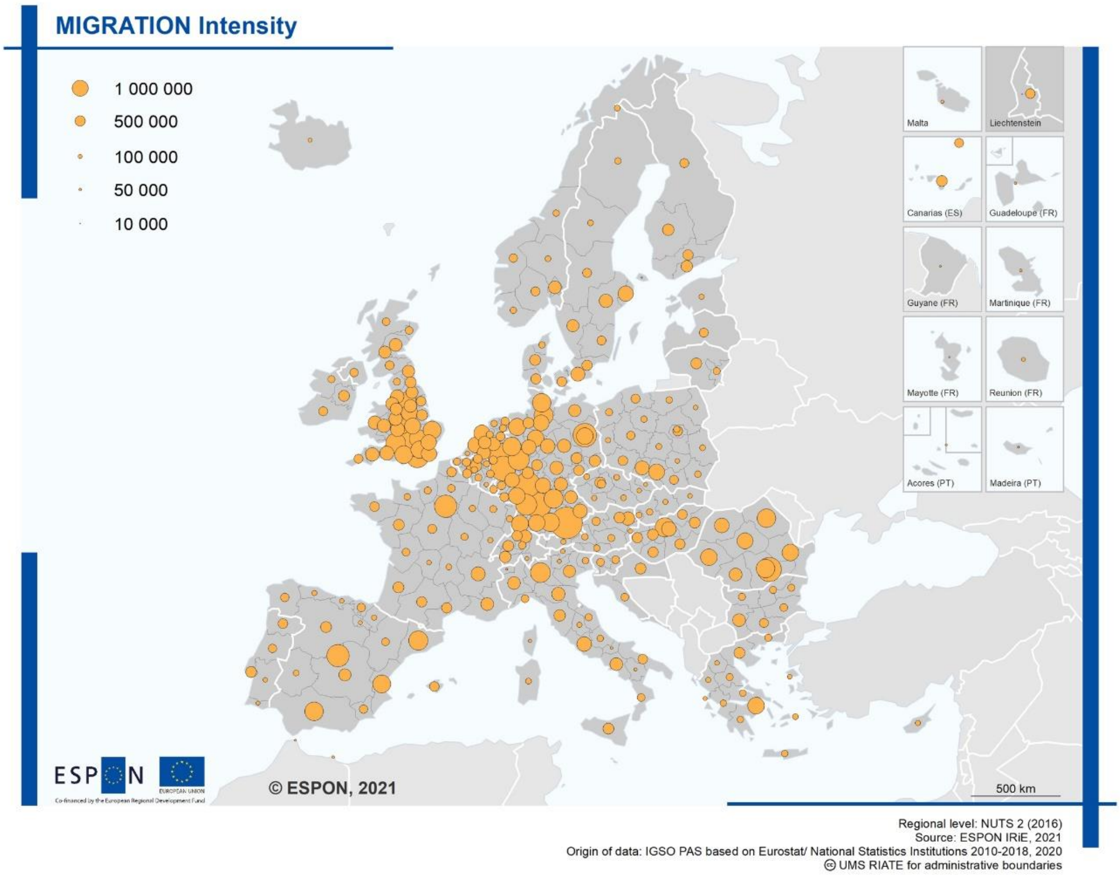


Analysis of the strongest migratory relations may also shed light on the internal spatial structure of the countries considered. Multidirectional flows between units in a definite country indirectly demonstrate polycentricity (one of the commonly used indicators; see ESPON 1.1.1 Final Report). At the same time, connections oriented around a single centre imply a monocentricity of spatio-functional structure. Our study confirms the polycentric character of the German, British (especially English), and Dutch settlement networks. The networks in Sweden and Spain, as well as in Romania, also turned out to be polycentric. On the other hand, migratory monocentricity was observed in France, Greece, and Bulgaria.

3.2.1.2 Intensity of migrations

The overall (absolute) intensity of migrations, presented in Fig. 3.14, was, as already indicated, partly conditioned by the magnitudes of the respective units in particular countries. Yet, in spite of this, the overall intensity ~~it~~ allowed us to identify certain essential territorial regularities. Intensity is expressed as the sum of migration inflows and outflows. It encompasses both domestic and international movements. Consequently, high values were observed in places where in the years 2010-2018 all such movements were significant. This applies to the countries concentrating incoming movements from abroad (to fill gaps in the labour market, strong “pull factors”, mainly in the form of relatively high wages), with a simultaneous high intensity of domestic migrations (including, in particular, movements of specialized personnel between centres, but also within metropolitan areas in the case of small-surface NUTS 2 units). Such a situation was observed primarily in Germany (mainly the western lands), Belgium, the Netherlands, and the UK (mainly England). The spatial pattern of Europe's migratory core so understood was related partly to the so-called Blue Banana, although the intensity of flows distinctly decreased in the southern direction, with values in northern Italy being perceptibly lower.

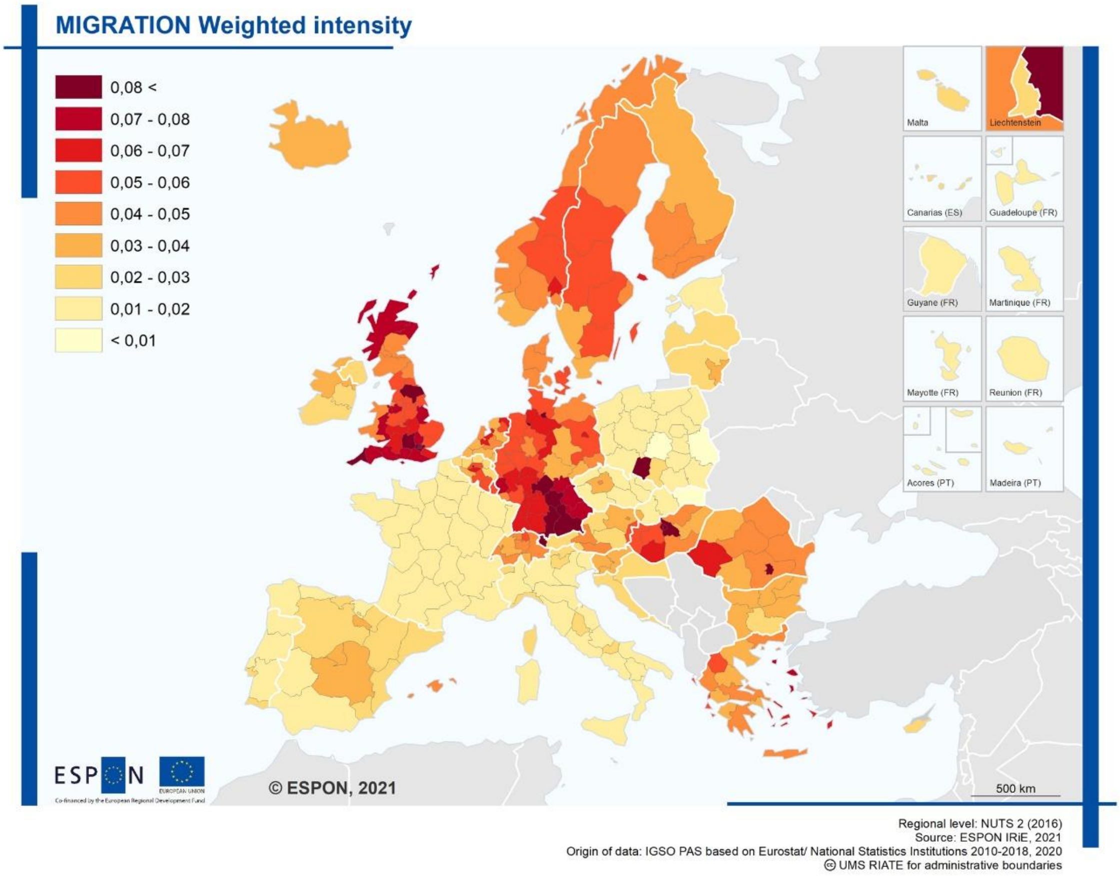
Figure 3.14: Intensity of migrations



We observed a characteristically low absolute intensity of migrations in France (except for Paris, and even though the NUTS 2 units are large). Areas surrounding the "migration core" featured diverse migration intensities. High values characterized areas of intensive outflow (and often also of numerous returns) as well as certain centres that domestic migrants find attractive (capitals and other economically strong regions). The first of these categories included, first of all, the Romanian regions, and also certain regions of southern Poland, while the second category encompassed the NUTS 2 units, which include such cities as Berlin, Madrid, Barcelona, Rome, Athens, Budapest, and Stockholm. It should be remembered in this connection that the high values for Spanish and Romanian units resulted also from the character of the NUTS 2 units in these countries. An apparently low intensity was observed in part of the peripheral regions, where a significant outflow (frequently abroad or to the capital of the country) was not accompanied by a registered inflow. These were the units that featured depopulation. One can classify into this category the regions of eastern Poland, Bulgaria, and Slovakia. Regions with persistently low migration intensity were, of course, also those that are thinly populated and geographically distant, such as northern Scandinavia and far-off territories (except for the relatively high intensity for Canary Islands).

The role of migrations in demographic development, and indirectly in a more generally conceived regional development, is illustrated by the indicator weighted by population (Fig. 3.15). This perspective eliminates in part the problem of the units' different magnitudes. It does not resolve the difficulty of capturing the more local migrations in these countries, where the units are smaller. In effect, the intensity of migrations in Germany and the UK was still overestimated for this reason (not fully comparable with the rest of Europe).

Figure 3.15: Migrations weighted by population



Particularly high values of the migration mobility indicator with respect to population were observed in the UK and in western Germany. It is characteristic that this largely concerned the areas with high population potential. The phenomenon might be explained at least partly by the magnitude of the NUTS 2 units. High values were observed also in all of Scandinavia (relatively large migration inflow coupled with low population density), Switzerland, Romania, and Bulgaria.

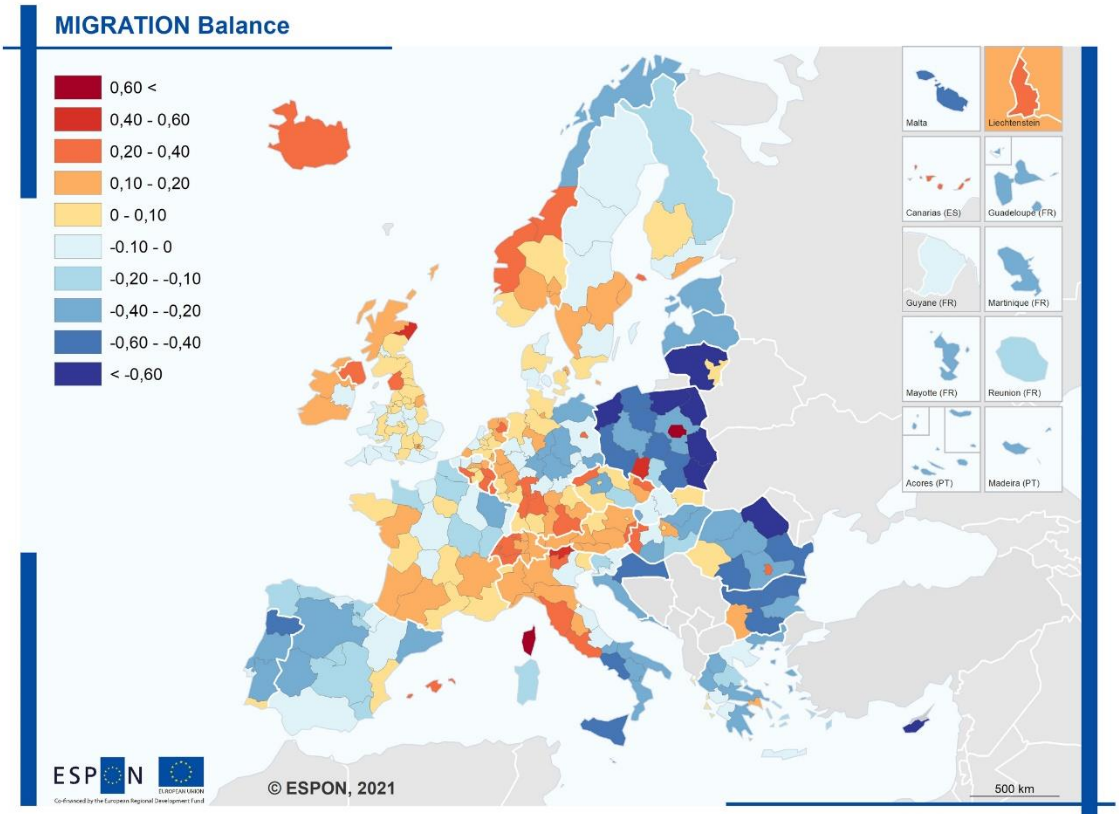
In the countries of southeastern Europe the indicator's high values might have been due to a large migration outflow with simultaneously low yet often decreasing populations. At the same time, the capitals of certain countries (Budapest, Bucharest) seemed to reflect a summation of outflow abroad with domestic inflow.

Migration intensity with respect to population was, simultaneously, low in such larger countries as France, Italy, and Poland. One explanation may be the present study's omission of migration inflow from outside the ESPON space (from North Africa to France, and from Ukraine to Poland). An exception was again provided by Poland's Opolskie region, with its strong bilateral migration relations with Germany.

3.2.1.3 Migration balance

The true attractiveness of regions is well represented by the indicator for net balance of migration inflows and outflows (see Fig. 3.16). A compact area with a clearly positive migration balance stretches from the Netherlands through western Germany, Austria, and Switzerland and down to the middle of Italy. Yet, at the same time, many countries are in this respect highly diversified internally. A characteristic spatial image was observed for the UK, where London, Scotland, and Northern Ireland feature a positive balance. At the same time, much of England and Wales was characterized by a negative balance. Geographic divisions known from earlier studies were observed in Italy (negative balance in the south and positive balance in the north and center) and in Scandinavia (dominance of outflows in far north). A different image was observed in Central-Eastern Europe, from the Baltic states and eastern lands of Germany down to Greece. This vast area is dominated by NUTS units with a negative or even deeply negative balance (especially eastern Poland, Lithuania, and eastern Romania). At the same time, capital regions displayed a clear domination of inflows over outflows. The instances of such "islands" of migration surpluses were the regions encompassing Warsaw (indicator value among the highest in Europe), Berlin, Budapest, Bucharest, Sofia, Vilnius, and Athens. A different situation, more advantageous in terms of the relation between inflows and outflows, was observed in some border-adjacent units in Central Europe, especially those located at the western borders of their respective countries. This phenomenon did not, however, appear in Poland, where the regions bordering Germany were characterized by a decidedly negative migration balance. Along with Warsaw, the positive balance was observed in Poland only in the previously mentioned region of Opole (strong relations to Germany, with domination of the return flow).

Figure 3.16: Migration balance



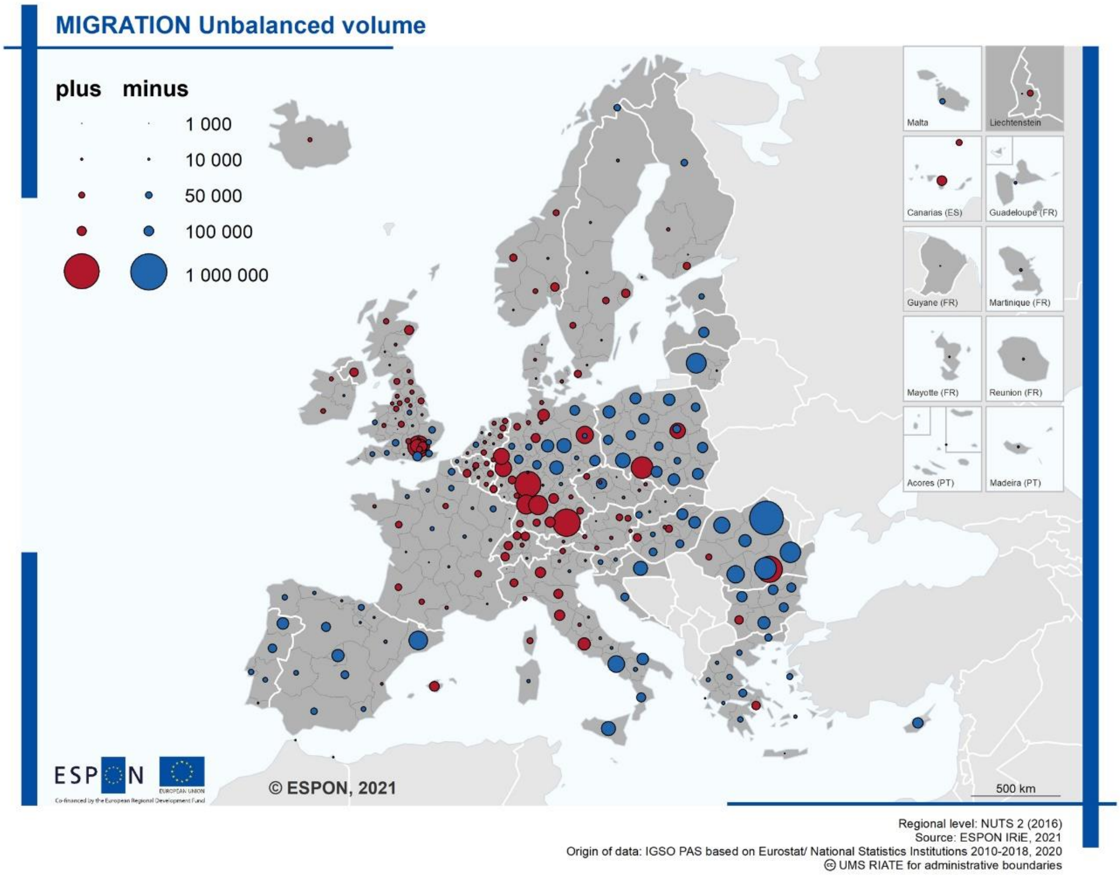
Regional level: NUTS 2 (2016)
Source: ESPON IRiE, 2021
Origin of data: IGSO PAS based on Eurostat/ National Statistics Institutions 2010-2018, 2020
© UMS RIATE for administrative boundaries

Distinctly positive migration balance values were also observed in the island regions of Spain (Baleares, Canary Islands) and of France (Corsica). There was, however, no analogy with the Italian (Sicily, Sardinia), Portuguese (Azores, Madeira), or Greek islands. A very clear dominance of outflows was noted for Croatia, Malta, and Cyprus as well. Regions featuring negative migration balance dominated very distinctly in Spain and Portugal (except for parts of the southern seacoast). Outside of the territories of these countries, which joined EU in 2004 and later, the most disadvantageous migration balance values in the period 2010-2018 were noted in western Spain, northern Portugal, southern Italy, northern Norway, the former East Germany (excluding Berlin), and northern France.

In Fig. 3.17 the magnitude of the negative or positive migration balance is presented in absolute terms. This makes it possible to assess migration-related deficits and surpluses, and hence also the degree to which they contribute to depopulation and population concentration respectively. There are numerous areas where the degree of imbalance (whether negative or positive) was altogether quite limited. This may have resulted from the changing situation in the successive years of the analysed period, 2010-2018. From the point of view of internal European migrations the population of these units was stable. The respective example was provided primarily by

France, but to a certain degree also by Sweden, Norway, Finland, Greece, Ireland, Austria, and Slovakia.

Figure 3.17: Absolute value of migration imbalance



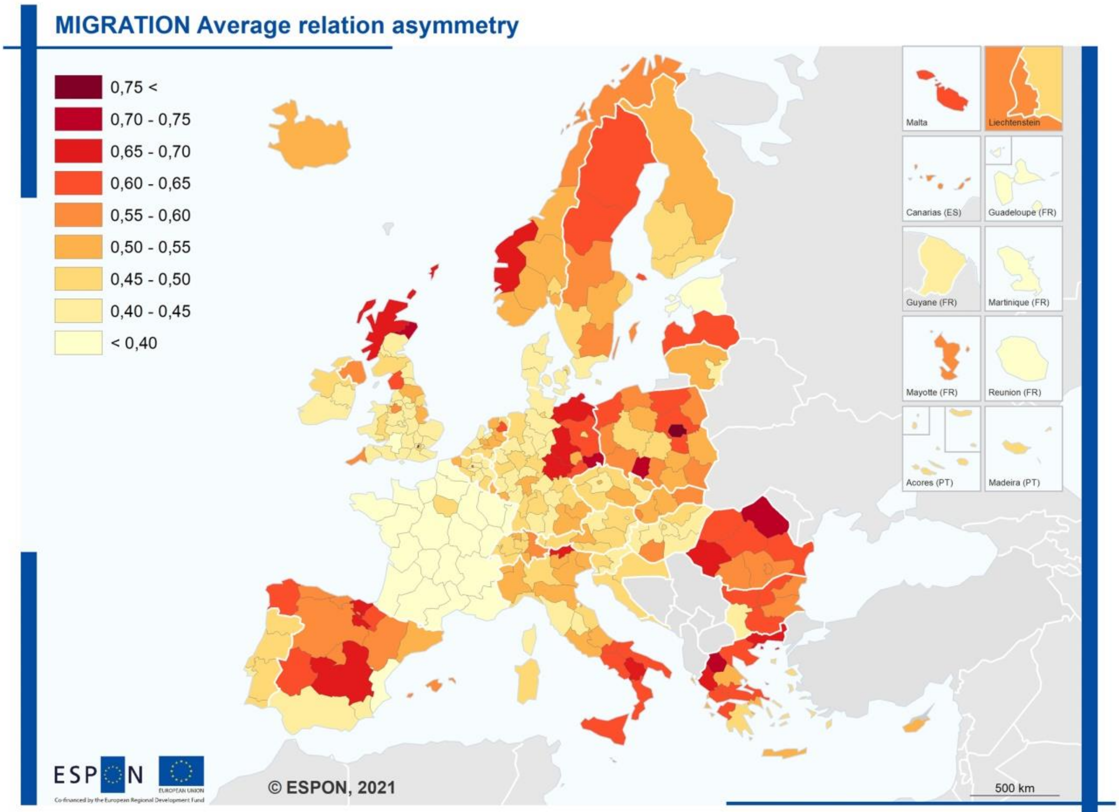
The areas that unambiguously concentrated absolute population numbers — as concerns the movements inside the ESPON space — were primarily the biggest cities of the western part of Germany (in particular Frankfurt am Mein, Stuttgart, Cologne, and Munich) as well as London. The cities that clearly advanced their demographic position in the settlement system of Europe in the years 2010-2018 were also Berlin, Warsaw, Bucharest, certain Italian cities (along the belt from Milan, through Bologna and Florence and down to Rome), as well as certain Scandinavian and Central European capitals. In comparing the biggest centres of contemporary migration-based population concentration it is important to take surrounding and/or neighbouring units into account. This helps clarify whether an inflow is not the effect of local movements (including intra-metropolitan ones), as appeared to be the case in Bucharest and, on a smaller scale, also in Prague. If these capitals formed a single unit with their surrounding regions, their joint migration balance would be negative.

In sum, the migration balance of 2010-2018 very distinctly strengthened in only a few of the concrete European MEGAs: Munich, Frankfurt, London, Berlin, and Warsaw. For many of the others interregional migration movements did not significantly change their position in the European settlement system. We can speak of the large centres' having undergone a weakening of their position (a significant and supra-local absolute net outflow) with respect to Madrid, Barcelona, and Naples. Peripheral zones (on the European scale) were characterized by large total outflow in the east and south of the continent, and a perceptible inflow in the north. Our analysis seems to confirm the proposition of an advancing concentration of demographic potential within the "Blue Banana", Scandinavia, and a few separate areas of Central-Eastern Europe.

The period analysed was, at the same time, one with a significant inflow of financial means into the framework of the cohesion policy, especially for the EU's eastern and southern peripheries. Despite significant improvement in infrastructure and quality of life, however, these means have not exerted an effective influence on the migration balance in most of their target regions.

We gleaned significant additional information from the average asymmetry of relations (Fig. 3.18). This indicator illustrates the degree of migratory imbalance with respect to other regions. It represents not the overall balance but the average imbalance of all migration relations. It is therefore possible for the overall magnitudes of inflows and outflows to yield a zero balance, but the indicator of asymmetry shows a strong imbalance for individual relations (in whichever direction). High values in this indicator may characterize regions that send out numerous migrants in definite directions, while simultaneously receiving incomers from other parts of the given country or from elsewhere in Europe. Asymmetry can, however, appear also in regions with generally low intensity of migration flows (like northern Scandinavia).

Figure 3.18: Average relations asymmetry



Regional level: NUTS 2 (2016)
Source: ESPON IRiE, 2021
Origin of data: IGSO PAS based on Eurostat/ National Statistics Institutions 2010-2018, 2020
© UMS RIATE for administrative boundaries

High levels of asymmetry were noted mainly within the peripheral areas of the ESPON space. In these areas, though, the asymmetry concerned both regions with a negative and regions with positive simple migration balance. The highest value of this indicator — on the European scale — was noted for Warsaw, and resulted from the multidirectional migration inflow from the territory of Poland, with a simultaneous and significant flow of outward foreign migrants. A similar situation, although on a smaller scale, was visible also in Bratislava, Prague, Budapest, and Bucharest. These cities can be considered migration “nodes” (common links in migratory movement chains). This role was not played by other capitals we’ve previously indicated as “islands” of positive migration balance (Berlin, Madrid, Athens). This may be the effect of different directions in their migratory inflow and their bigger share of pan-European inter-metropolitan movements (highly skilled employees).

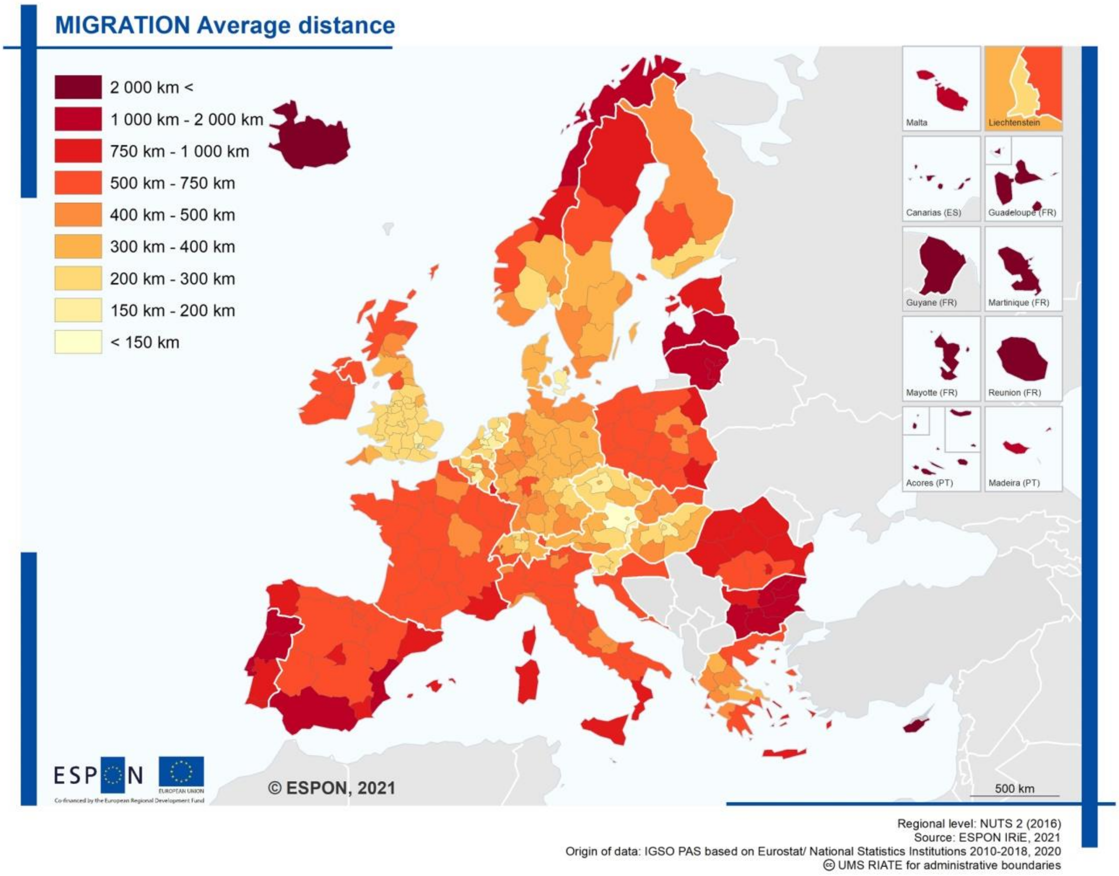
High values of asymmetry were noted also in Romania, Greece, Bulgaria, Spain, and Scotland. There was a characteristic differentiation in this respect with Germany. The western lands featured low asymmetry, but the situation was quite different in the former territory of East Germany (with exception of Berlin and Leipzig). This most likely signifies a multidirectional migration outflow to various units in the western part of the country. An analogous explanation probably holds for the high values of the indicator, observed in southern Italy.

The centrally located core areas of the European Union and most British regions were characterized by a lower average asymmetry of migration relations. Internal movements between well-developed regions (flows of skilled labour) were relatively balanced out. Migrations from the peripheries constituted only a part of total flows (in contrast to the peripheral areas themselves, where the outflow dominated both on the European scale and domestically) and did not much influence the level of the indicator. The lowest asymmetry was observed for the French regions.

3.2.1.4 Migration distance

Distance is one of the basic elements of spatial relations (as described by the gravity model, including numerous studies devoted to migrations). Its significance was visible in external migrations — lack of political barriers — but also in international ones, especially when the labour market was not subject to excessive formal constraints. In our study we calculated a simple indicator for every NUTS 2 region of the average migration distance (inflow and outflow for the period of 2010-2018, see Fig. 3.19).

Figure 3.19: Average migration distance



The image acquired agrees to a great extent with our working hypothesis on the necessity of longer-distance migrations from and to geographically peripheral areas. The highest value of this indicator was noted for the most peripheral areas on the European scale, especially where the role of foreign migrations was pronounced (Bulgaria, Romania, eastern Poland, Portugal) and where the country's geographical dimensions compelled long domestic migrations (Italy, Norway). Large average migration distances appeared also in southern Spain and southern France. These may be the effect of both the magnitude of the country (internal movements at long distances) and the inward movement of foreign migrants (e.g. persons in post-productive age from Northern Europe). Long migration distances were noted in a natural manner for peripheral countries whose entire area constitutes a single NUTS 2 unit (Iceland, Latvia, Estonia, Cyprus) and for overseas territories.

Shorter average migration distances were observed in countries located more centrally and/or characterized by small NUTS 2 units. (Accounting for a part of the local migrations shortens the average migration distance in a natural manner.) This concerned in particular Germany, the Netherlands, Belgium, Austria, Switzerland, Czechia, Hungary, and the UK. The indicator took exceptionally low levels in the regions of certain metropolises, such as London and Vienna.

Interesting things took place in the surroundings of several of the peripheral countries' capitals: specifically when a metropolis constituted a relatively small NUTS 2 unit, surrounded by one or more other statistical units. The core of this group of regions had a longer average migration distance, while its surroundings featured distinctly shorter distances. This indicates the strong local influence of these MEGAs, with a simultaneous concentration in them of far-off migration connections (including international ones). Neighbouring units were characterized by short average migration distance, since these were mainly migrations to the closest strong metropolis. Such a metropolis tended to be attractive insofar as it took advantage of the migration resources not only of its direct surroundings, but also of more distant regions. Moreover, the centre would participate in international migratory exchange (movements of skilled labour). Such a city in Central-Eastern Europe could serve also as a link in a migration chain. Domestic migrants, attracted to the metropolis, would sometimes undertake further migration abroad. This was the situation in Berlin, Warsaw, Prague, Bucharest, and Budapest but also in Madrid, Helsinki, and Stockholm. The effect of the aforementioned factors in the shaping of the situation was, however, different for individual cases. There were significant migration flows from the northern fringes of Sweden and Finland. In Berlin, though, the development of service and cultural functions were more important in attracting migrants, even from quite distant regions of Europe.

The indicator for average migration distance showed, as well, core-periphery system for some of the countries considered. This was clearly visible in the Scandinavian countries, Switzerland, and the UK as well as in Hungary and Slovakia.

A complement to the information on the significance of distance in the distribution of migration flows in Europe is the indicator of dependence upon distance (Fig. 3.20). It shows the intensity of influence exerted by distance on the magnitude of migrations. Its value ranges from 0 to 1, where 1 denotes full conditioning of the magnitude of migration by the distance between the

NUTS 2 units in question. This influence can take place in regions with short or long average distance of migration relations.

The indicator's spatial distribution is different than in the case of our other metrics. The greatest significance of distance was in the Alpine countries (Austria, Switzerland), as well as in Wales and western England. High values were noted in Scandinavia, Spain, the Netherlands, Slovakia, eastern Germany, western and southern France, and central Italy. We can safely say that distance played a characteristic role in transitory areas between the cores and the peripheries of Europe that were neither the main recipients nor the emitters of international migration flows. This proposition was partly negated by the high positions of Austria and especially of the UK. These two cases might be explained by intense domestic movements. In the countries of Central-Eastern Europe the indicator took on higher values in their western parts (Czechia, Poland, Slovakia, Hungary), which may be associated with a concentration of migrations over shorter movements, either to domestic metropolises or to nearby Germany and Austria. At the same time, the eastern areas of these countries displayed a clearly weaker dependence of migratory intensity on distance. In these areas foreign migrations took place often over very long distances, these movements being accompanied by much shorter movements to neighbouring units.

The indicator's lowest values (lack of dependence upon distance) were observed both in the core of the European Union (western Germany, northern and eastern France — inflow from many regions, situated at highly differentiated distances) and within the eastern periphery (eastern Poland, Lithuania, Latvia, Romania, and Bulgaria; analogously: outflow to numerous regions, situated in various parts of Europe).

Figure 3.20: Dependence upon distance

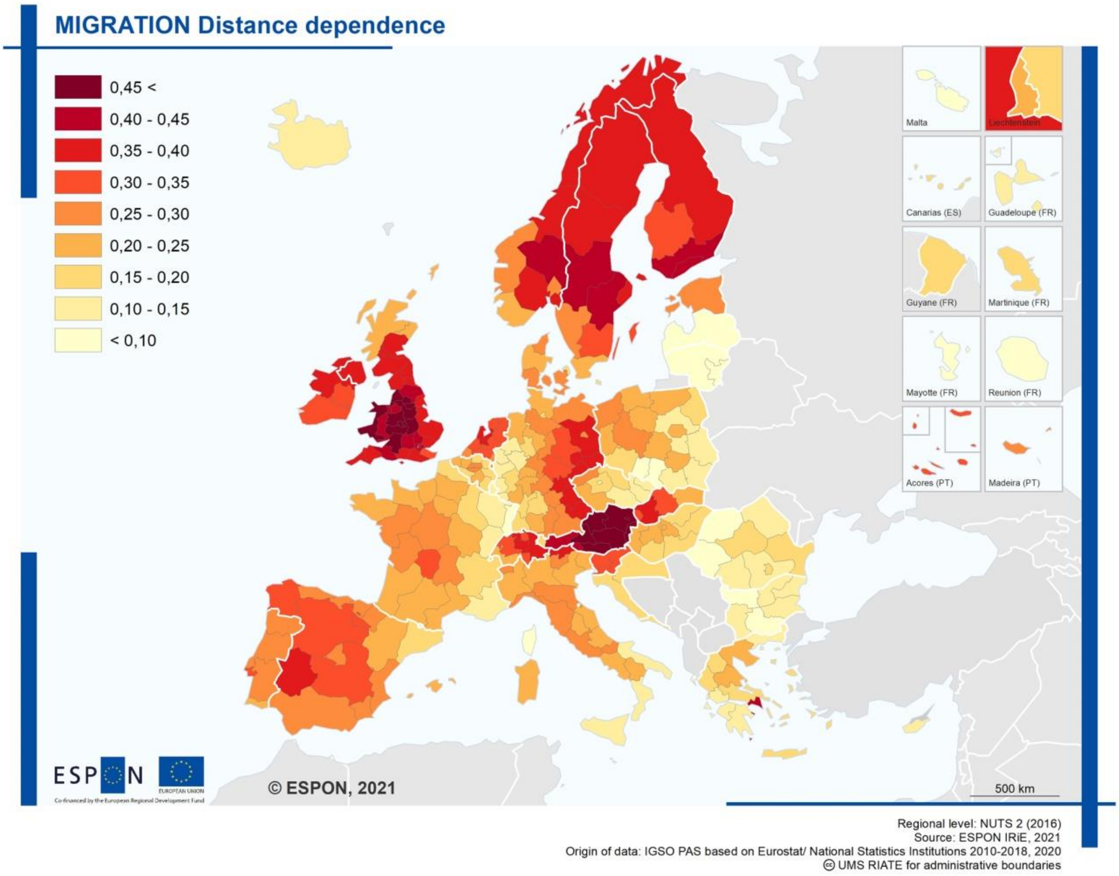
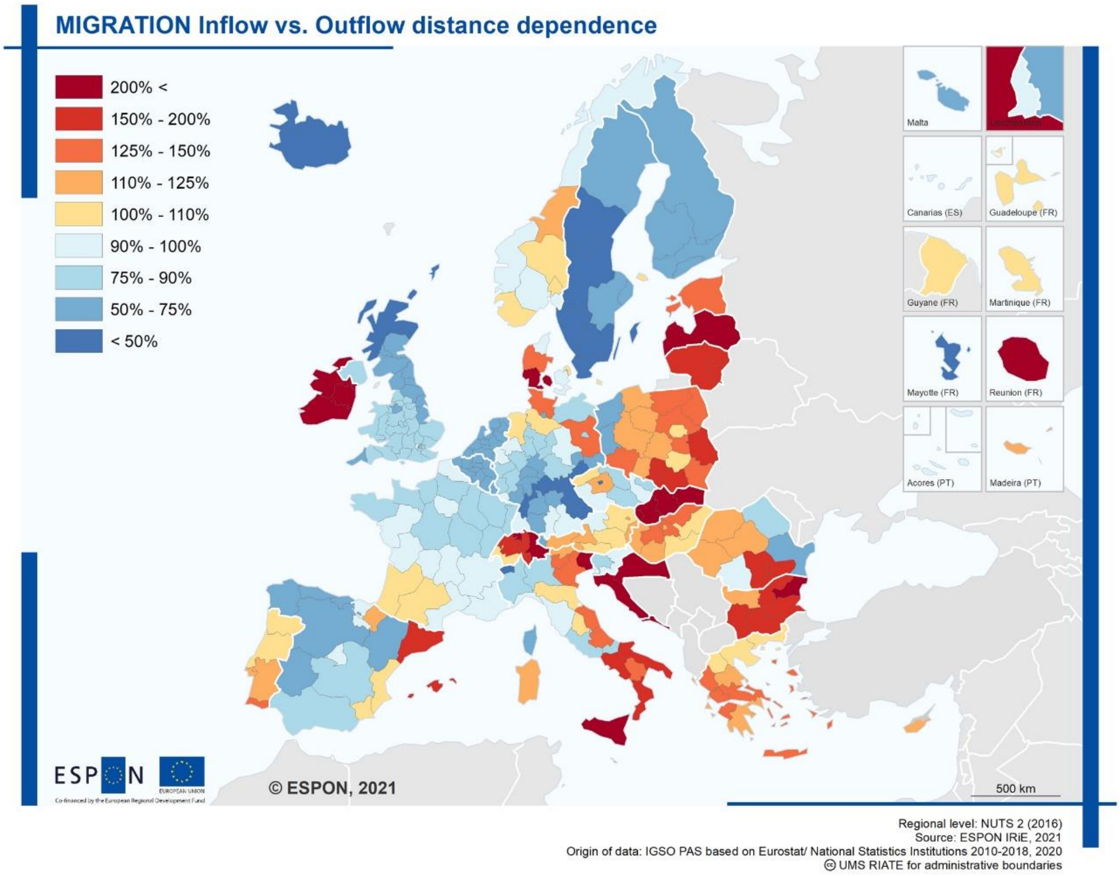


Figure 3.21 presents the indicator's mutual relations of dependence upon distance for migration inflows and outflows. This allows us to determine which of the two basic directions of the relation has the stronger geographic conditioning, and which depends upon other factors. The division was distinct in the eastern and western parts of the ESPON space. The former showed strong dominance of the relative significance of distance to inflow rather than outflow. Emigration from these areas was most obviously conditioned by other factors. At the same time, immigration was smaller in scale and more regional in nature. We can put into the same group certain regions on the peripheries of Europe other than the eastern ones (southern Italy, Portugal, Ireland, individual NUTS 2 in Norway).

Figure 3.21: Dependence upon distance. Relation of indicators for inflows and outflows

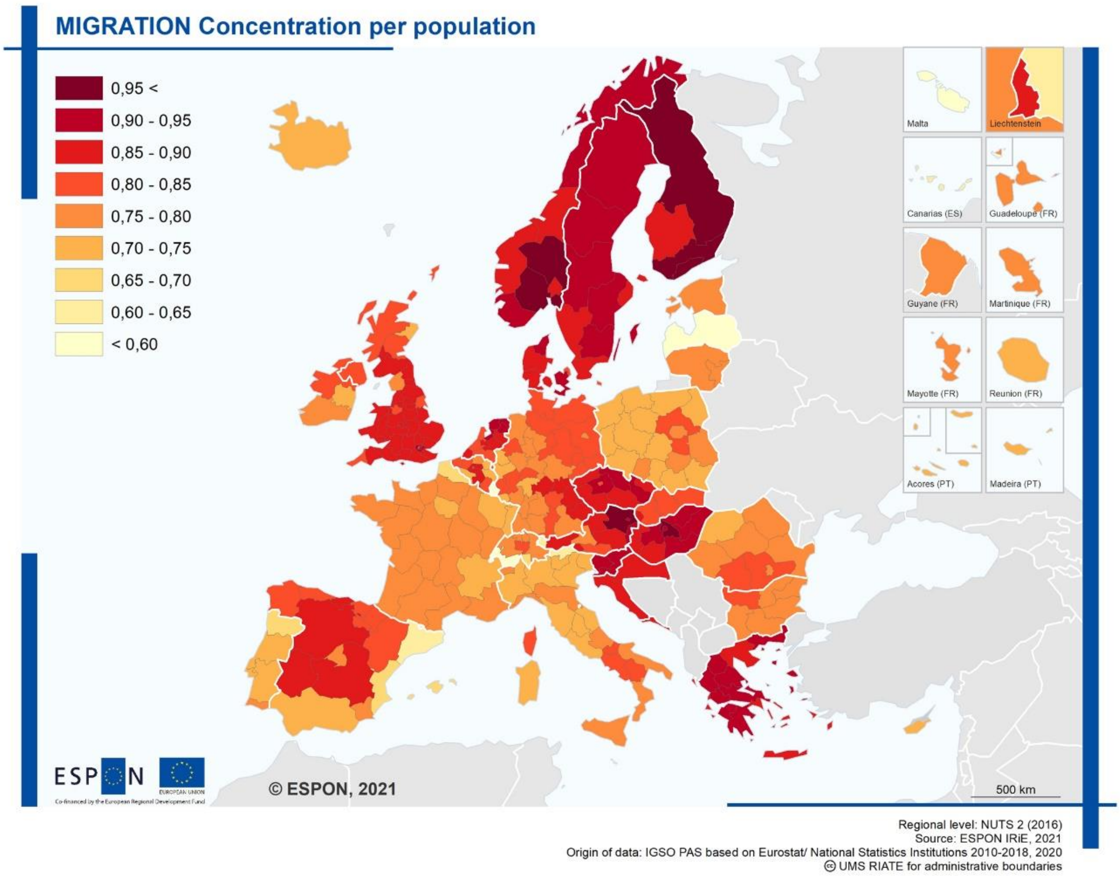


On the other hand, most of the countries of western Europe (including Spain, the UK, and Sweden) were characterized by the reverse situation. The role of distance appeared there mainly to have concerned migration outflow. Inflows originated from multiple, sometimes quite dispersed locations. Outflow was often constituted by domestic movements, conditioned by distance. In the countries of Central-Eastern Europe the abovementioned stronger dependence of inflow (rather than outflow) with regard to distance was less perceptible (or not observed at all) in the capitals, such as Berlin, Warsaw, Budapest, and Prague.

3.2.1.5 Concentration of migration relations

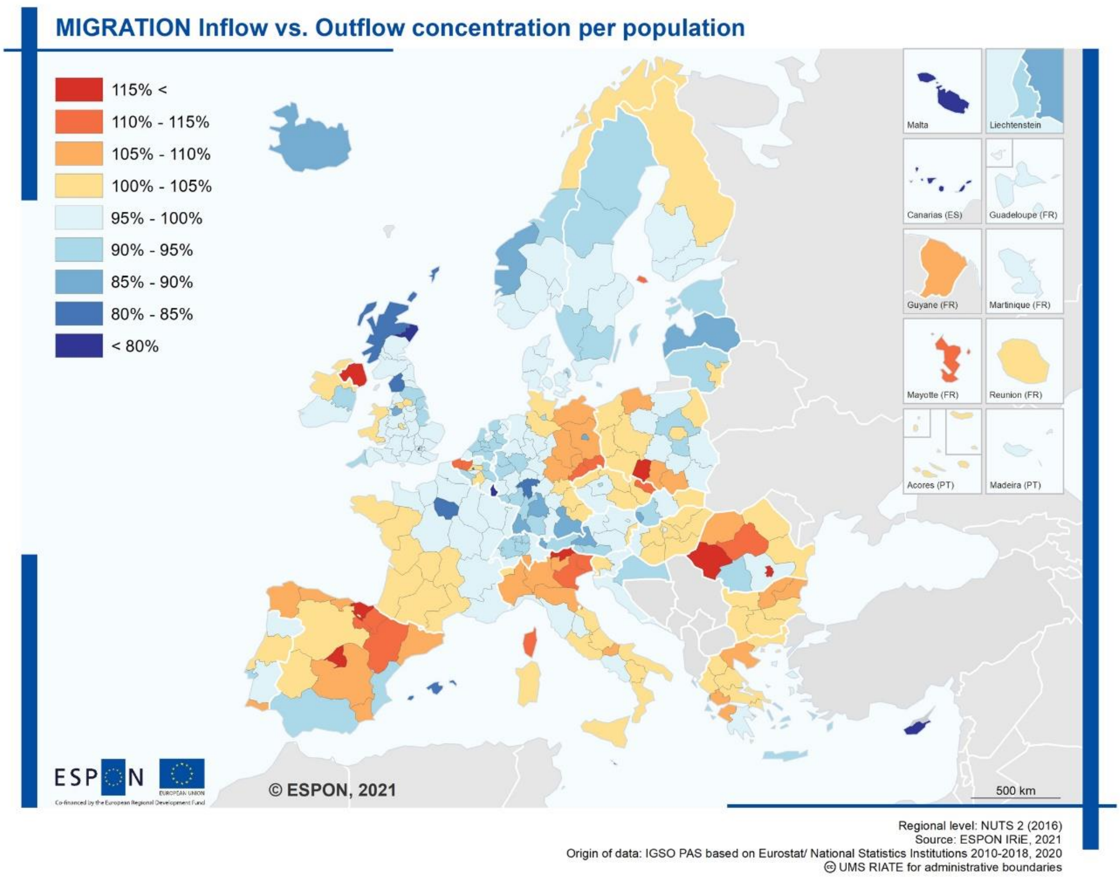
The concentration of the migration relations was assessed in the study with respect to area (territory) and to population. In each case we compared the concentration of migration outflows and inflows. Given the small differences between the two approaches, it was decided to present only the distribution of the indicator values of concentration with respect to population. Fig. 3.22 shows the general bilateral concentration. This illustration uncovers, first of all, domestic patterns, featuring high concentrations of internal migrations and few metropolises, that attract migrants. Examples were Hungary, Austria, Czechia, the Netherlands, Norway, Sweden, Denmark, and Finland. Lower values of the indicator were observed, as a rule, in bigger countries, both in Western Europe and in Central-Eastern Europe. This may be due to more polycentric settlement patterns, and also to dispersed international migrations (whether inflows or outflows). In peripheral countries higher levels of indicator values were observed usually around capital units and other metropolises (but not in their very cores). This is the evidence for the common draining of migrants from the direct hinterlands of the metropolises. These appeared irrespective of the degree of economic development, but they seemed to be more intensive in countries with lower GDP per capita than the EU average. High concentration was characteristic of British regions, which ought to be associated with the generally high percentage share of domestic movements with respect to total migrations. As in the previous aspects of analysis, internal differentiation was clearly visible in Germany — the level of concentration being higher in the eastern lands, as well as in northern and eastern Bavaria.

Figure 3.22: Concentration of migration flows with respect to population



The concentration of general migration flows did not reflect the full reality, especially in regions with more balanced flows. That is why Fig. 3.23 shows the ratio of the indicators of concentration of inflow and outflow, revealing which of the two basic directions remained more concentrated in the sense of regions-partners (with consideration of their respective population potentials). This approach made apparent new regional divisions inside the countries considered. Higher concentration on the side of migration inflows (the origins of the migrants showing relatively little spatial diversity) was observed in a couple of compact areas in Europe. One of these encompassed eastern Germany, western and southern Poland, and northern Czechia; another northern Italy, northern and central Spain, and Hungary, along with northern Romania. In some cases (Central Europe) the dominant concentration on the side of inflow can be associated with the generally low, mostly neighbourhood immigration, with simultaneous quite significant and dispersed emigration (e.g. from eastern Germany to numerous NUTS 2 regions in the western part of the country). Other compact areas observed may be the consequence of long-term migration traditions (as in northern Italy, supplied with immigrants from the south of the country). Higher concentration on the inflow side (with respect to surrounding areas) showed up also in the capitals of certain peripheral countries (Warsaw, Bucharest). This was not the case in Berlin, Prague, and Budapest, however. These differences might result from the pattern of administrative boundaries. They may, however, also reflect differences in the positions in migration chains. In Warsaw and Bucharest migration inflows originated mainly from domestic sources. In Prague and Berlin an important part of the inflow comprised migrants from farther areas, some of them abroad.

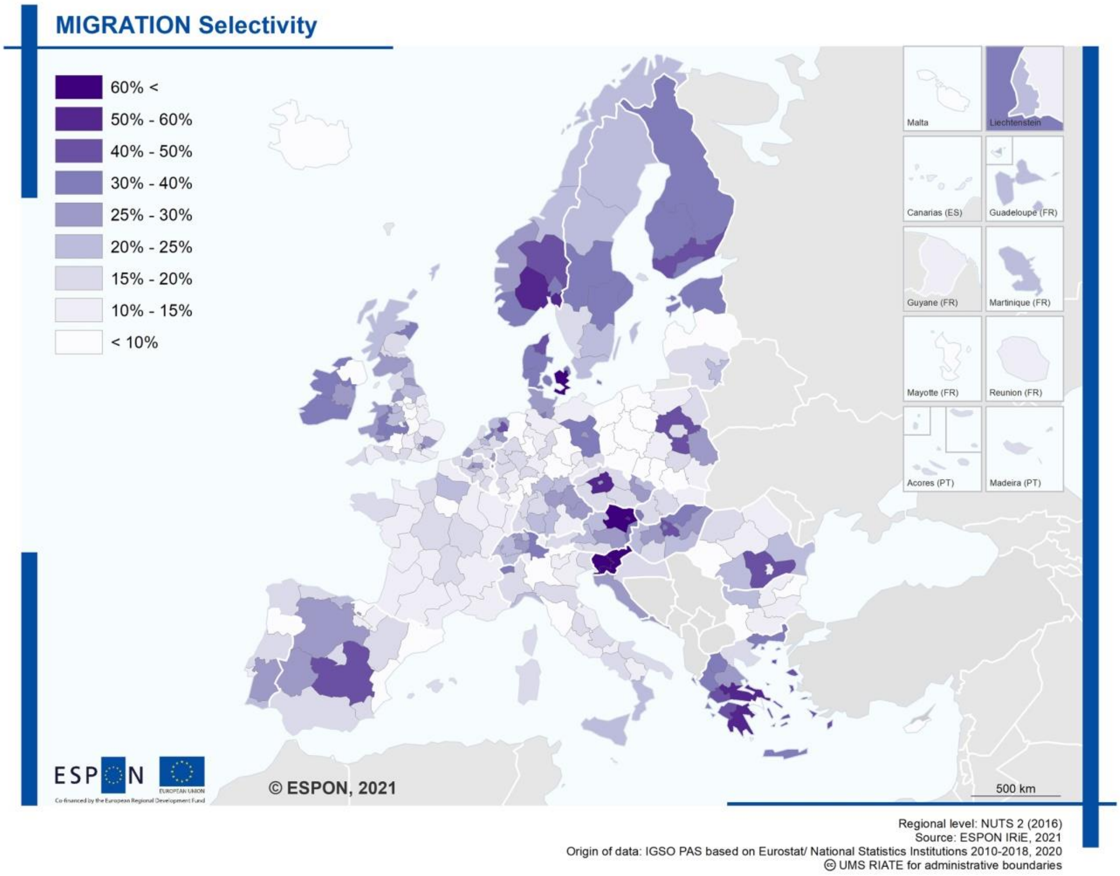
Figure 3.23: Concentration of migrations with respect to population. Ratio of inflow to outflow indicators



The distinctly dominant concentration on the outflow (rather than inflow) side occurred in the immigration regions. We observed this in western Germany, the UK, northern France, Belgium, the Netherlands, Austria, and Switzerland. In these countries outflow was directed to a limited number of the most attractive destinations, while the migration inflow came from many regions, including those in the peripheral countries of the EU. A similar mechanism may be decisive for the bigger relative dispersion of the inflow in residentially attractive southern Spain. Besides, a higher relative concentration of outflow was observed also in parts of certain countries of Central-Eastern Europe (e.g. eastern Poland, Slovakia, southern Romania) as well as in the Baltic countries and on the Greek islands. The reasons can be sought in the concentrated outflow to the neighbouring capital centre (Warsaw, Bucharest, Athens).

A complement to the indicators of concentration is provided by the assessment of selectivity and the external influence, referring in each case to the percentage share of the strongest relation. Fig. 3.24 shows the selectivity of migrations, presenting the share of the strongest relation from the side of the migrants' region of origin. High values in this indicator imply strong territorial orientation of migration flows. The highest values might be interpreted as measuring the extent to which a given region constitutes the key migration source for some other region.

Figure 3.24: Migration selectivity (from the side of the region of origin)

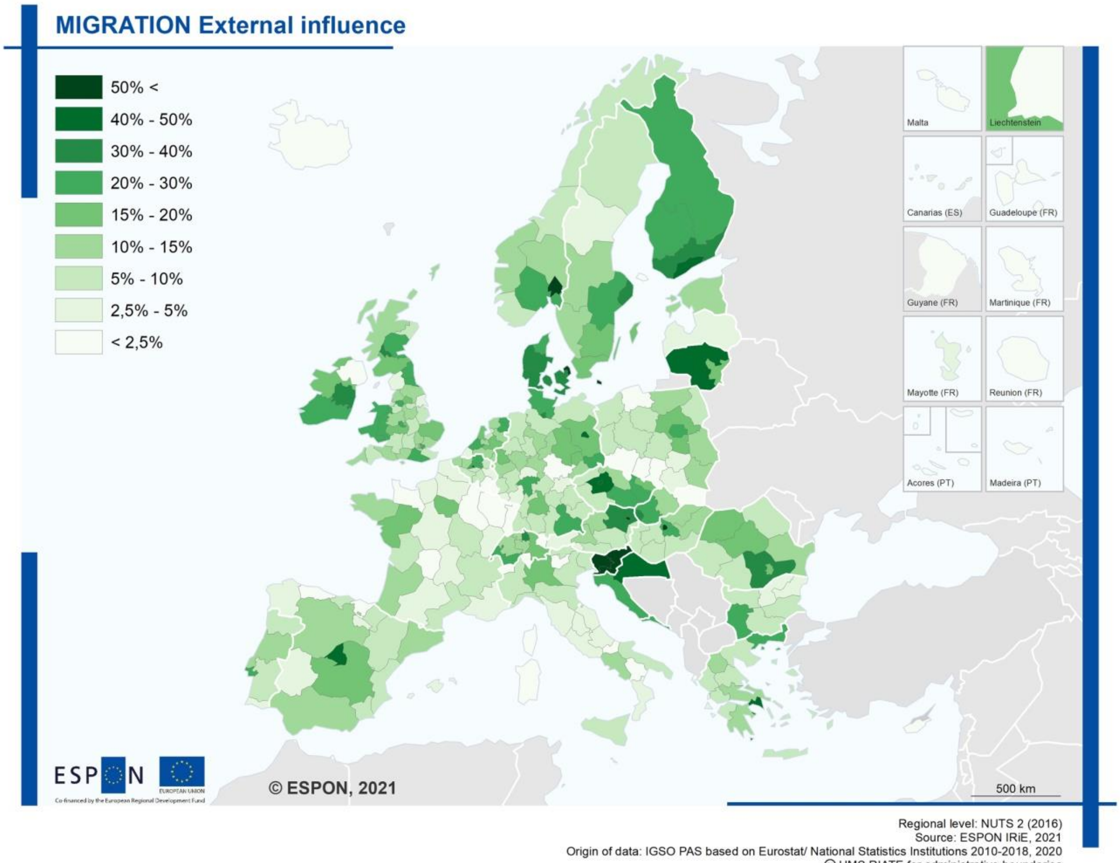


This indicator took on its highest values for certain peripheral areas of the ESPON space, usually near metropolises (frequently country capitals) but outside of their cores. This was, again, visible in the surroundings of Berlin, Prague, Vienna, Budapest, and Bucharest, as well as Athens and Madrid. High values were also observed for the Scandinavian countries, Ireland, and Slovenia. There were cases where the indicator's value exceeded 50%. These values were clearly lower in the core of the European Union, including the western part of Germany, the entirety of France, and the entirety of Italy.

An essential observation is the diversification of indicator values in the group of regions with a strong negative migration balance. Some of them featured a geographically multidirectional outflow (in the regional perspective — e.g. southern Italy, western Poland, western Romania and Bulgaria, eastern Germany — except for Brandenburg), without the domination of a single "receiver" region. For the others domination was observed often from the side of the local metropolis. Such was the case for eastern Poland and eastern Romania, as well as Hungary, Scotland, and Wales.

An inverse of the indicator of selectivity is the indicator of external influence on migration (see Fig. 3.25), which illustrates the percentage share of the strongest outflow relation from the side of the receiving region. High values in this indicator mark a dependence of regional labour markets on the inflow of migrants from a specific other unit. This occurred in the cores of large metropolises, such as London, Paris, Lisbon, Copenhagen, Berlin, Madrid, Warsaw, Budapest, Vienna, and Oslo. The migration hinterland of these centres most often comprised their surrounding areas. Yet the percentage shares here were usually lower than for the indicator of selectivity. Inflows to large cities are the largest from surrounding areas, but as a rule many immigrants in metropolitan areas also have other roots.

Figure 3.25: External-influence indicator



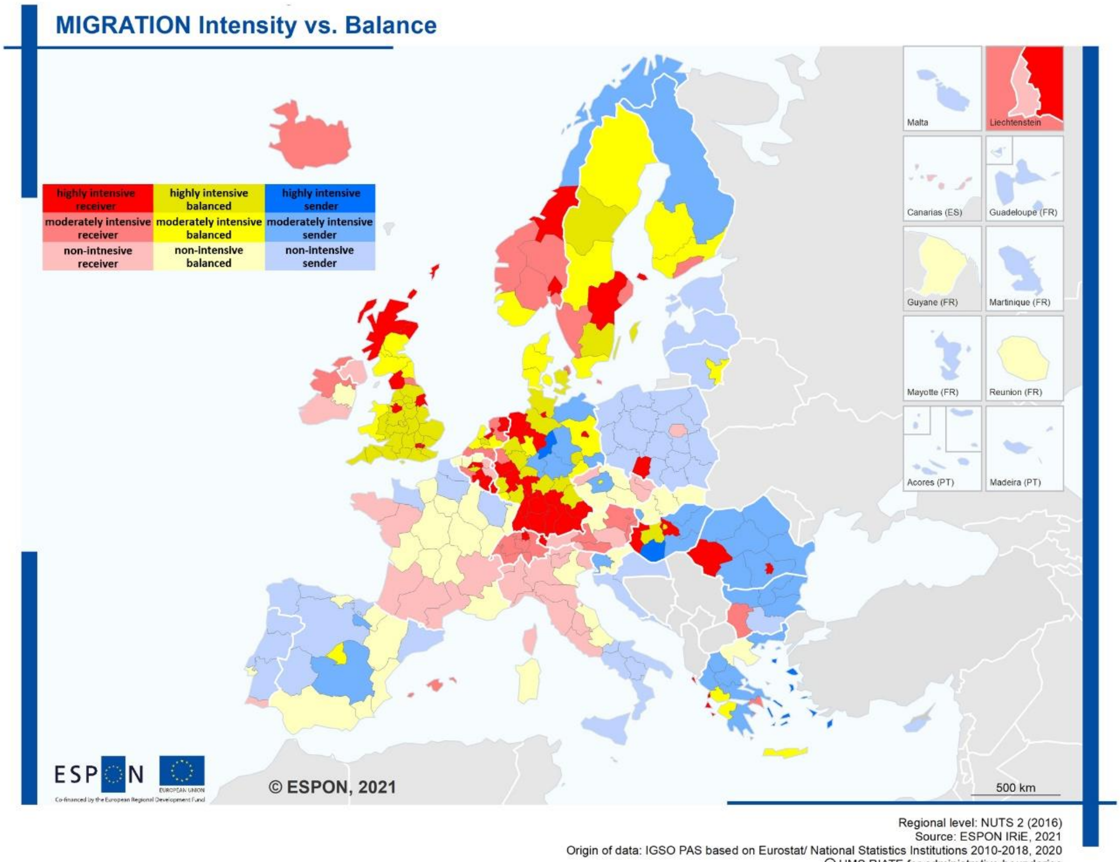
Regional level: NUTS 2 (2016)
Source: ESPON IRiE, 2021
Origin of data: IGSO PAS based on Eurostat/ National Statistics Institutions 2010-2018, 2020
© UMS RIATE for administrative boundaries

In many countries the territorial distribution of this indicator's values had a mosaic-like pattern. High values did appear, sometimes in peripheral areas (on both the country and the European scale). This implies the existence of barely identifiable migration connections between various pairs of regions. Often, but not always, these were neighbouring regions. Very low values of this indicator suggest a broad spectrum of immigration sources.

3.2.2 Typologies

The studies performed enable us to propose simple typologies, based in each case on two variables. Fig. 3.26 presents a division based on migration intensity and migration balance. We have established nine basic types of regions, classified into units of intensive, moderate, and weak migration flows and, simultaneously, receiver, sender, and balanced units. The image obtained makes it possible to indicate the most important outflow and inflow areas in Europe. A strong or moderate domination of outflow coupled with a high intensity of migrations was observed first of all in Romania, Bulgaria, and Greece. These areas were followed by eastern Hungary, most of the units of eastern Germany, northern Finland and Norway, and central Spain. The greatest domination of outflow among the regions with most intensive migrations was observed in south-western Hungary and along the former border between East and West Germany. Many other areas of distinctly emigrational character (negative migration balance) had a relatively weak general intensity of migrations. This concerned the most of the units in Poland, southern Italy, western Spain, and Portugal.

Figure 3.26: Typology of regions based on migration intensity and balance



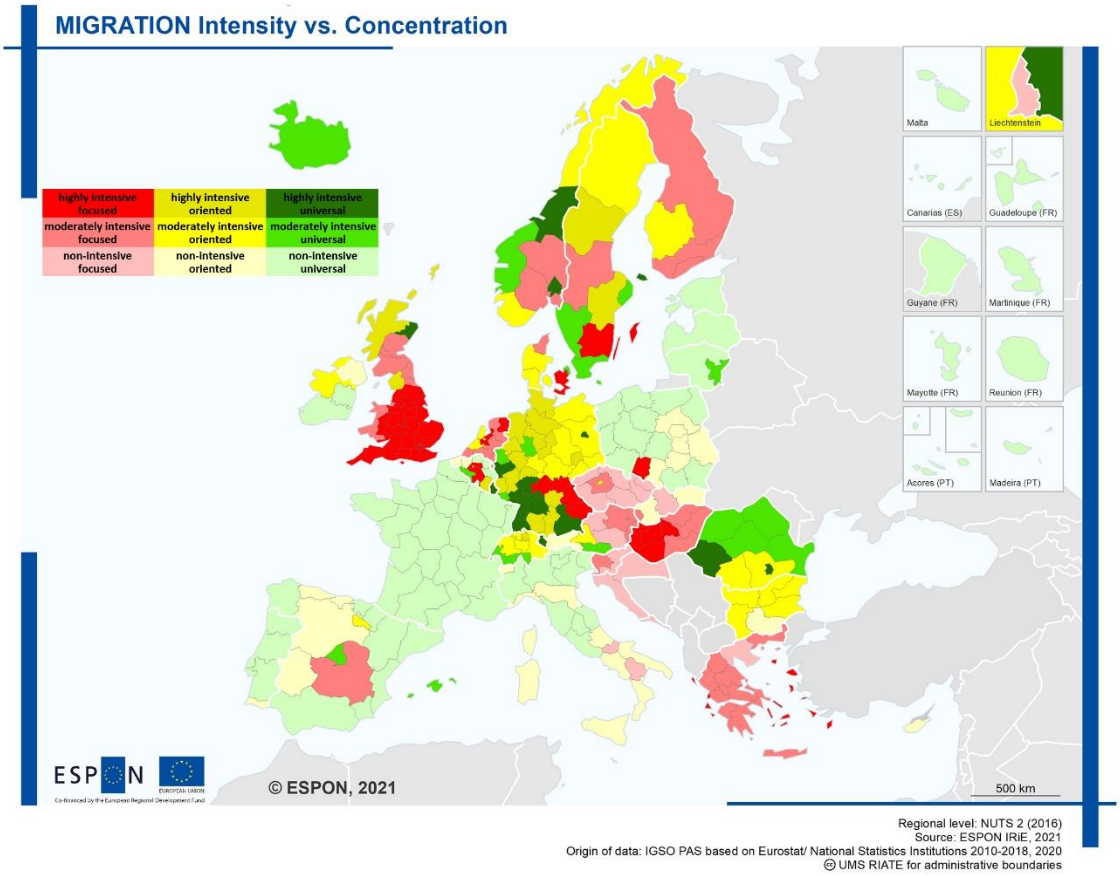
Regional level: NUTS 2 (2016)
Source: ESPON IRiE, 2021
Origin of data: IGSO PAS based on Eurostat/ National Statistics Institutions 2010-2018, 2020
© UMS RIATE for administrative boundaries

Regions characterised by strong, imbalanced migration inflows constituted a compact area primarily in Germany, where they formed by a belt stretching from Bavaria through Baden-Württemberg and Hessen up to Rhineland-Northern Westphalia. Similar situations existed in Lower Saxony, middle Sweden (including the neighbourhood of Stockholm), Norway, Switzerland, Ireland, and northern Scotland. One can put into the same group certain metropolises in Germany (Hamburg, Berlin), the UK (London, Liverpool), and Central-Eastern Europe (Bucharest, suburban zone of Budapest). Intensive and at the same time balanced migrations encompassed, firstly, almost the whole of England (except for the metropolises just mentioned), while vast areas with a positive balance of migration flows in northern Italy and southern France showed a distinctly lower overall intensity of migrations.

Almost no regions of intense and equilibrated migration balance appeared in Southern or Central-Eastern Europe.

The subsequent typology of the NUTS 2 units is based on the joint consideration of intensity of migration flows (bilateral) and degree of concentration (Fig. 3.27). High intensity of migration movements with simultaneous concentration of these movements was observed primarily in the UK (England), western Bavaria, and western Hungary, as well as in singular units elsewhere in Europe (including, in particular, the aforementioned Opolskie region in Poland and certain Belgian and Dutch regions). Parts of other outflow regions (central Spain, Greece) were characterized by strong geographic orientation while featuring flows of medium intensity. Others (like southern Italy, northern Spain, and eastern Poland) displayed less intensive migrations and moderate spatial concentration.

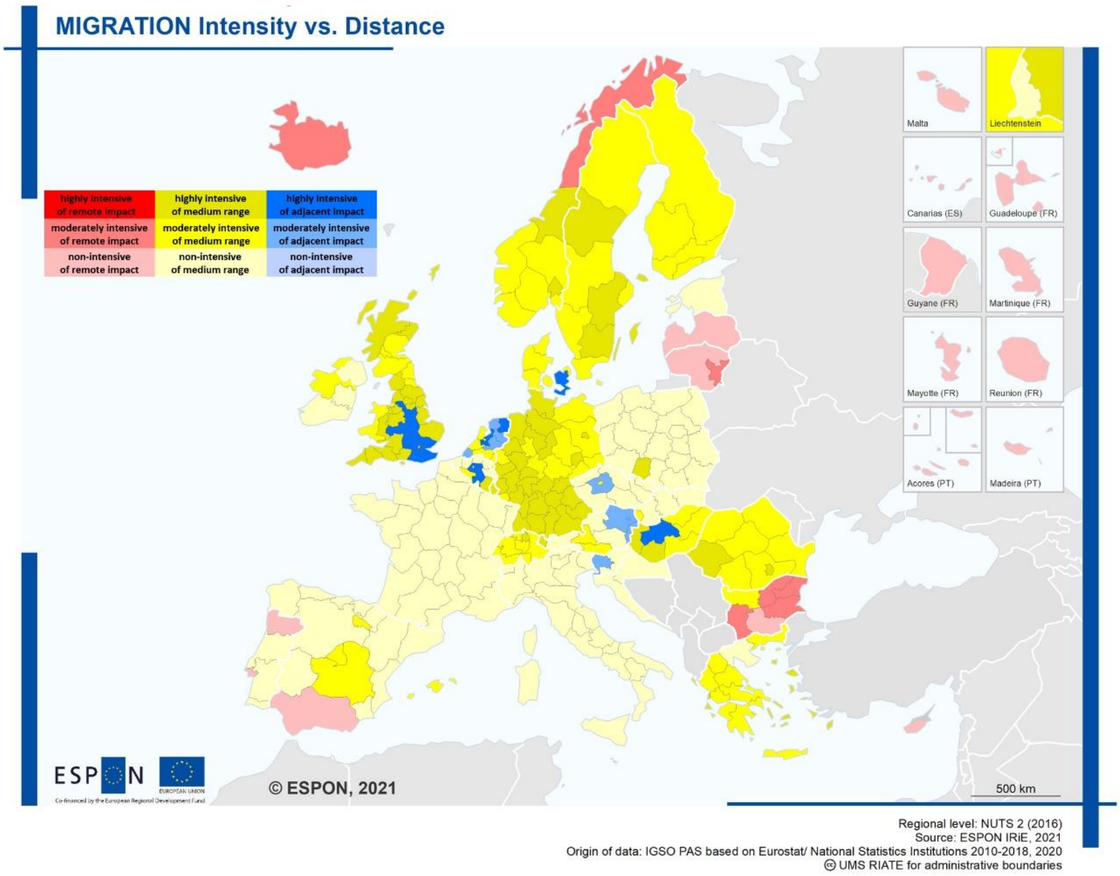
Figure 3.27: Typology of regions based on intensity and concentration of migrations



Parts of regions with high intensity of flows featured low concentration as well. This concerned both migrant-receiving and migrant-dispatching units. The former comprised the regions of southern Germany, middle Norway, southern Sweden, and certain metropolitan centres (Berlin, Oslo, Bucharest). Chief among the latter cases were northern and western Romania. Vast areas in France and Poland, and also southern France, northern Italy, and Portugal, were characterized by low intensity and, at the same time, significant territorial dispersion of migration flows.

By typology, based on intensity and average distance of migrations (Fig. 3.28), the dominant regions featured flows of both moderate distance and relatively high intensity. This is understandable, given the significance of domestic migrations in some of the countries considered. Western Germany and eastern as well as western England especially featured medium migration distances and high migration intensity. Of similar type were individual Swedish, Norwegian, and Romanian units. It was notable that relatively few regions were of extreme types. High intensity coupled with short migration distance was observed primarily in central England (along with the major part of Greater London), the Netherlands, southern Belgium (Wallonia), eastern Denmark (Zeeland with Copenhagen), central Hungary (with Budapest), and Czechia (particularly the area surrounding Prague). It was characteristic that this type did not include German units, even though they are (like the British ones) territorially small and were characterized by a high intensity of flows. This implies that the more distant migrations of western Germany (including both the international ones and those originating from the eastern lands of Germany) were more significant than those of the UK. Finally, it is understandable that high- or medium-intensity migrations coupled with significant distances of flows were observed for distant regions (northern Norway, Bulgaria, Iceland).

Figure 3.28: Typology of regions based on migration intensity and distance



Our overall assessment of the three typologies (in connection with the analysis of the basic indicators) indicates the separate character of some of the territorial clusters in terms of flow intensity, balance, concentration, and distance of migration. A plausible working hypothesis would be that these are, in particular:

- Areas of intensive migration with a distinctly positive balance, relatively low concentration, and differentiated migration distances. This cluster includes southern Germany (Bavaria).

- Areas of intensive migrations with an equilibrated balance, strong territorial concentration, and short migration distances. This cluster includes central England (together with London) as well as central and western Hungary (together with Budapest).
- Areas of intensive migration with a distinctly negative balance of flows, differentiated concentration, and rather long migration distances. This cluster forms a belt stretching from eastern Hungary through Romania and Bulgaria to Greece.
- Areas with moderate intensity of migrations and moderately negative migration balance, a noticeable level of territorial concentration, and an average distance of migrations. Examples of such compact areas were southern Italy, northern Spain, Portugal, and eastern Poland.

In addition, a couple of groups of geographically dispersed units share certain features. Chief among these are:

- Certain (not all) overseas territories and geographically distant territories (including Cyprus, Malta, northern Norway) with moderate or, in some cases, high migration intensity, negative balance of flows, low concentration and long migration distances;
- Residentially attractive areas with high intensity of migrations, a clearly positive balance of flows, low concentration, and long or medium migration distances. This type included southern Spain, southern Portugal, the Canary Islands, Baleares, and Corsica;
- The metropolises of Central-Eastern Europe, with at least medium intensity of migrations, positive balance of flows, low concentration, and differentiated migration distances. Here the examples were Berlin, Warsaw, Bucharest, Sofia, and Athens.

3.2.3 Explanatory factors: drivers and barriers

The moving forces of migrations become increasingly complex, changing over time (Wesselbaum and Aburn, 2019). Generally, following the literature, we can classify these factors from a global perspective into three categories: socio-economic, political, and climate-related. The last category has been the subject of increasing interest. In addition, there are numerous studies indicating that migrants tend to settle where migrants of the same ethnic origin or the same country of birth have settled before (for the significance of networks and diasporas, see Bredtmann et al., 2020). We are, of course, considering differences in regional development in our interpretation of population migrations (Zhu et al., 2021).

The results are subdivided into ten model specifications; the first and second models concern the main gravity variables and the GDP of both pairs of regions: i.e. the population of the regions of origin and destination, the distance between them, and either GDP per capita delayed by one year (model 1) or household revenues delayed by one year (model 2). Since both models yield similar results as far as the income component is concerned, our successive model specifications, 3 to 11, account solely for GDP per capita (except for model 6). Models 3 to 10 include all the push and pull factors (e.g. innovativeness of the economy, stability of the labour market, "language community"), along with factors related to the location of regions (border adjacency) and geographical aspects (metropolitan areas, country regions from before the EU extension of 2004) or economic aspects (locations within areas with freedom of movements of people, the euro zone). Model 10 contains three zero-one variables, concerning the period that has elapsed since the introduction of the freedom of movement of people between particular pairs of countries (and regions) with distinctions for long, medium, and short periods of time from the removal of barriers to the flow of people. On the other hand, model 1 simply accounts for the addition of the variables describing the labour market (employment) in source regions as a potential stimulant to migration flows. The authors of the report have established also that the GDP of receiving countries is a very important and encouraging factor for migrants. The GDP of source countries is also a statistically significant pushing factor for migration flows, in corroboration of numerous earlier studies (see Wesselbaum and Aburn, 2019; Mayda, 2010; Manzoor et al., 2021). Yet, in view of the correlation of GDP in general with the variable in relative terms, we decided for the last models to interpret the latter as referring to the wealth of the regions while omitting the aspect of their sheer magnitude.

Each of the model sets was elaborated in several variants:

- in totality, i.e. for all R2R migration flows (Table 3.2);
- with origins in border regions and non-border regions (Table 3.3);
- with destinations in border regions and non-border regions (Table 3.3);
- with origins in regions of new and old EU member states (Table 3.4);
- with destinations in regions of new and old EU member states (Table 3.4);
- with origins in regions with and without metropolitan areas (Table 3.5);
- with destinations in regions with and without metropolitan areas (Table 3.5);
- with migrations between regions within countries (domestic; Table 3.6);
- with international migrations between European regions (Table 3.7).

In analysing the general model for each of the model specifications presented here for all the flows in their totality we have concluded that distance is a determining factor with the minus sign, this being a clear confirmation of the gravity model (see Manzoor et al., 2021). This model confirmed also the significance of associations between the magnitude of migration flows and income levels as well as the wealth of the receiving (positive sign) and the source (negative sign) regions. Along with this, the size of the population in the regions of origin and destination is statistically meaningful for the respective models and amplifies the volume of migration flows between the regions. Another variable of interest is the distance between the primary languages of the home country and of the region of destination. An affinity of languages (the very same language or same group of languages) essentially enhances the magnitude of migration flows (see Bredtmann et al., 2020; Manzoor et al., 2021). Our study took into account the environmental factor (namely, the relation of mean annual temperature and precipitation between the regions of origin and destination), a variable often used to investigate migrations on the global scale (e.g. Backhaus et al., 2015). As could be expected, in Europe this factor has a lower influence: precipitation turned out to be statistically significant only in model 4 (for $p<0.01$), while temperature came out as significant in two specifications (models 6 and 9, both with a minus sign).

One might expect the labour market and the perspective of employment, as well as the development of essential sectors of the economy, would exert a visible influence on migrations (see Qi et al., 2021). In the general model, however, unemployment in the regions of origin did not seem to exert a sufficient push (except in model specification 11), but a bad labour market in the region of (potential) destination did constitute a discouraging factor (see model specifications 4 through 11). We can therefore conclude from model specification 11 that for source regions employment and its stability, especially in key labour groups (persons in the mobile age, women), have higher significance. And, of course, high levels of employment among young persons in (potential) regions of origin discourages outmigration.

We expected, a priori, that regions of origin where a high percentage of the population has received higher education would produce more intensive migration flows. The general model confirmed this hypothesis (specifications 5 through 11). As for political factors, membership in the Schengen zone constituted an essential strengthening factor for migration flows, but for membership in the euro zone the dependence is not very obvious (see specifications 4 through 11). In turn, belonging to a country that joined the EU after 2004 (in terms of the region of origin) significantly amplifies the magnitude of migration flows (specification 6). Freedom of movement is crucial for the ESPON space. From the results for the specification 10, then, we can conclude that the opening up of labour markets in particular regions stimulated migrations. The shorter the time since the opening up, the stronger the influence (novelty effect). In labour markets that have been mutually open for a long time we observed an inverse relation.

Migrations were also strengthened by metropolitan centres, acting as both regions of destination (engines of growth) and regions of origin (magnitude of the centres, flow of people having definite competences). (In this connection see specification 8.) A high degree of urbanisation in the place of origin constituted an essential, statistically significant factor, limiting the respective migration flows (see specification 9). For the model of total flows as well as for all models taking into account flows in both directions (border/non-border regions, metropolitan/non-metropolitan regions, and old/new EU Member States) domestic migration flows were a statistically significant variable for increasing migration flows.

The model specified for inflows to/outflows from border regions shows, additionally, the qualitative factor associated with the economy of the border regions and their frequent perception as peripheral areas. High outlays in R&D within these areas significantly limited outmigration (see specifications 5 through 11). This correlation was significant also for the models specified for outflows from new EU member countries.

Contrary to the general model, the model for inflows to non-border regions showed that the magnitude of migration flows was positively influenced by bad labour markets (high unemployment) in regions of origin.

The model for inflows to new member countries of the EU and non-metropolitan regions omits the significance of economic factors — GDP per capita (specification 1) and disposable income of households (specification 3) — for the destination region. Here, an important stimulus to the flow of migrants was internal movements, frequently undertaken for non-economic reasons (e.g. family ties, attractive locality, accessibility of amenities).

Quite similarly, the economic criteria of GDP and income at the disposal of households lost their significance in the model for the outflows from regions of the “old” EU member countries (for regions of both origin and destination). Moreover, there was no statistical significance to the influence of employment on the labour market. We can therefore suppose that non-economic factors might affect these kinds of flows. The same situation was observed in models for outflows from metropolitan regions.

The model for internal (domestic) migrations did not show statistical significance for GDP per capita as a stimulant to migration flows (this being of importance for international migrations). Yet income in regions of origin and destination was significant (the most attractive regions inside a country had higher incomes in the population). Likewise, the remaining economic factors (such as employment levels except among the young, unemployment levels, and the structure of the economy) were not statistically significant in the model. What turned out to be significant for domestic migrations was belonging to metropolitan regions (in regions of both origin and destination, inter-metropolitan migrations of staff, students, etc.).

The model of international migrations was altogether similar in terms of its statistics to the general model. In the general model unemployment in regions of origin did not constitute an effective push, but a bad labour market in regions of destination did constitute a discouraging factor for migrating (see specifications 4 through 11). In contrast to the general model for total flows, for international migration also the level of unemployment in the origin region was statistically significant in shaping migration flows (with a negative sign). We can therefore conclude that the model for international flows does not necessarily deal with levelling out inequalities in the labour market (oversupply) and flows from areas with more difficult labour markets to areas with labour shortages. It is more of a competition between income and job quality (required qualifications, wages, and their purchasing value).

Table 3.2: Estimation results of the econometric model for explanatory factors of migration flows between European regions in 2010-2018 – entire dataset

	dependent variable: $Migration\ flow_{ijt}$									
VARIABLES	(1)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
popul_o	0.816***	0.829***	0.731***	0.851***	0.824***	0.993***	0.687***	0.974***	0.850***	0.075
popul_d	0.787***	0.779***	0.667***	0.725***	0.652***	0.677***	0.631***	0.828***	0.728***	0.693***
dist	-1.511***	-1.605***	-1.034***	-0.959***	-0.932***	-0.954***	-0.960***	-0.996***	-0.943***	-0.989***
lag_gdp_pc_o	-0.313***		-0.298***	-0.433***		-0.301***	-0.434***	-0.566***	-0.272***	-0.345***
lag_gdp_pc_d	0.164***		0.276***	0.382***		0.530***	0.358***	0.233***	0.410***	0.361***
pop_den_rel			0.019	0.037**	0.127***	0.055**	0.042***	0.013	0.041**	0.052***
domestic			2.593***	2.456***	2.138***	2.346***	2.497***	2.417***	2.314***	2.466***
language			-0.195***	-0.118***	-0.027	0.060*	-0.114***	-0.191***	0.111***	-0.118***
outer_o			0.724***	0.754***	0.723***	0.649***	0.775***	1.031***	0.819***	0.537***
outer_d			1.516***	1.507***	1.190***	1.373***	1.599***	1.588***	1.511***	1.426***
island_o			0.278***	0.244***	0.177*	0.172*	0.307***	0.288***	0.255***	0.376***
island_d			0.318***	0.146*	0.319***	0.148*	0.149**	0.168*	0.177**	0.198***
euro_rel			0.273***	0.071*		0.138***	0.029	0.128***	-0.526***	0.108***
precip_rel			0.131***	0.047	0.061*	0.046	0.069**	-0.067*	0.043	0.042
temp_rel			0.007	-0.000	-0.042***	-0.000	0.001	-0.021***	-0.005	0.005
unempl_o			0.045	0.066*		0.018	0.047	0.074**	0.083**	0.218***
unempl_d			-0.140***	-0.209***		-0.176***	-0.237***	-0.237***	-0.218***	-0.235***
schen_rel			0.421***	0.477***		0.699***	0.561***	0.264***	0.595***	0.403***
lag_disp_inc_o		-0.734***								
lag_disp_inc_d		0.487***								
h_edu_o				0.381***	0.396***	0.675***	0.281***	0.780***	0.441***	0.356***
rd_exp_o				-0.049*	-0.036*	-0.207***	-0.041	-0.048	-0.058*	-0.078***
new_eu_o					0.454***					
agr_sh_o						-0.102***				
metro_o							0.383***			
metro_d							0.232***			
urban_o								-0.276***		
urban_d								-0.033		
lab_mar_long									-0.192***	
lab_mark_med									0.880***	
lab_mark_short									1.305***	
emp_y_o										-0.323***
emp_sen_o										0.109**
emp_w_o										1.061***
emp_mobil_o										0.805***
Constant	-9.18***	-5.78***	-9.75***	-13.62***	-12.51***	-13.96***	-9.88***	-18.67***	-13.32***	-10.37***
Observations	665,428	535,178	598,870	440,163	509,320	361,116	440,163	283,290	440,163	414,496
pseudo R-	0.387	0.402	0.518	0.492	0.480	0.471	0.483	0.592	0.487	0.499

Robust standard errors for significance tests; *** p<0.01, ** p<0.05, * p<0.1; the PPML estimation results.

Source: Own elaboration.

Table 3.3: Estimation results of the econometric model of migration for border and non-border regions in 2010-2018

	dependent variable: $Migration\ flow_{ijt}$															
	Outflows – border regions				Inflows – border regions				Outflows – non-border regions				Inflows – non-border regions			
VARIABLES	(1)	(3)	(5)	(10)	(1)	(3)	(5)	(10)	(1)	(3)	(5)	(10)	(1)	(3)	(5)	(10)
popul o	0.884***	0.866***	1.108***	1.141***	0.788***	0.789***	0.908***	0.952***	0.764***	0.783***	0.641***	0.572***	0.786***	0.860***	0.996***	0.989***
popul d	0.778***	0.762***	0.599***	0.577***	0.745***	0.735***	0.596***	0.597***	0.785***	0.812***	0.843***	0.868***	0.752***	0.749***	0.762***	0.752***
dist	-1.583***	-1.579***	-0.892***	-0.870***	-1.509***	-1.523***	-0.816***	-0.804***	-1.489***	-1.647***	-1.049***	-1.055***	-1.486***	-1.647***	-1.034***	-1.035***
gdp pc o	-0.508***		-0.280***	-0.113	-0.290***		-0.191*	-0.138	-0.231***		-0.242**	-0.282***	-0.428***		-0.292***	-0.048
gdp pc d	0.417***		0.571***	0.524***	0.275***		0.460***	0.472***	-0.024		0.216***	0.455***	0.080		0.385***	0.345***
pop den rel			0.145***	0.177***			0.045	0.051			-0.039*	-0.059***			0.104***	0.120***
domestic			2.343***	2.350***			2.393***	2.413***			2.472***	2.126***			2.422***	2.205***
language			0.096***	0.213***			0.146***	0.239***			-0.208***	0.137**			-0.238***	-0.007
outer o											0.786***	0.838***			0.686***	0.809***
outer d			1.048***	0.963***							1.827***	1.962***			1.329***	1.381***
island o			0.135	0.242			0.354	0.494**			0.302***	0.271**			0.251	0.261
island d			0.345***	0.417***			-0.105	0.311*			0.033	0.022			0.356***	0.357***
euro rel			0.051	-0.432***			0.162**	-0.398***			0.164***	-0.675***			0.038	-0.513***
precip rel			0.001	0.013			0.061	0.067			-0.013	0.007			0.079	0.061
temp rel			0.006	-0.003			0.005	0.001			-0.036***	-0.029***			0.013	-0.001
unempl o			-0.062	-0.091**			-0.067	-0.138**			0.228***	0.258***			0.148***	0.214***
unempl d			-0.259***	-0.280***			-0.351***	-0.353***			-0.094*	-0.050			-0.207***	-0.235***
schen rel			0.491***	0.436***			0.878***	0.565***			0.241***	0.674***			0.352***	0.576***
disp inc o		-0.905***				-0.595***				-0.658***				-1.119***		
disp inc d		0.692***				0.458***				0.170				0.609***		
h edu o			0.534***	0.611***			0.631***	0.689***			0.448***	0.511***			0.315***	0.289***
rd exp o			-0.226***	-0.246***			-0.181***	-0.216***			0.072*	0.095**			-0.107***	-0.102**
new eu o																
agr sh o																
metro o																

metro d																
urban o																
urban d																
lab mar long				0.112*				0.314***				-0.638***				-0.391***
lab mark med				0.937***				1.126***				0.812***				0.709***
lab mark short				1.383***				1.655***				1.432***				0.954***
emp y o																
emp sen o																
emp w o																
emp mobil o																
Constant	-9.52***	-6.67***	-13.84***	-14.03***	-7.83***	-6.31***	-11.99***	-12.81***	-8.82***	-2.86*	-13.16***	-12.03***	-9.01***	-2.90*	-14.74***	-13.61***
Observations	357,354	332,605	248,319	248,319	191,298	206,157	137,052	137,052	308,074	202,573	191,844	191,844	308,095	202,630	196,493	196,493
R-squared	0.378	0.369	0.522	0.514	0.338	0.324	0.468	0.470	0.484	0.479	0.583	0.580	0.472	0.498	0.555	0.542

Robust standard errors for significance tests; *** p<0.01, ** p<0.05, * p<0.1; the PPML estimation results.

Source: Own elaboration.

Table 3.4: Estimation results of the econometric model of migration for new and old EU member-state regions in 2010-2018

	dependent variable: *Migration flow*$_{ijt}$															
	Outflows – New EU member states regions				Inflows – New EU member states regions				Outflows – Old EU member states regions				Inflows – Old EU member states			
VARIABLES	(1)	(3)	(5)	(10)	(1)	(3)	(5)	(10)	(1)	(3)	(5)	(10)	(1)	(3)	(5)	(10)
popul o	1.024***	1.140***	0.907***	0.890***	0.937***	0.930***	0.909***	0.948***	0.820***	0.821***	0.656***	0.655***	0.828***	0.850***	0.903***	0.874***
popul d	0.864***	0.861***	0.855***	0.839***	0.772***	0.769***	0.277***	0.220**	0.801***	0.806***	0.727***	0.727***	0.810***	0.810***	0.843***	0.824***
dist	-1.714***	-1.781***	-1.146***	-1.124***	-1.701***	-1.710***	-1.092***	-1.070***	-1.554***	-1.640***	-0.987***	-0.973***	-1.534***	-1.635***	-0.992***	-0.976***
gdp pc o	-0.977***		-0.474***	-0.399**	-0.026		0.967***	0.775***	-0.330***		0.106	0.055	-0.434***		-0.780***	-0.093
gdp pc d	0.732***		1.283***	1.251***	-0.058		-0.299***	-0.149	-0.091**		-0.153***	0.241***	0.031		0.510***	0.427***
pop den rel			-0.148***	-0.131***			0.207***	0.168***			0.106***	0.077***			0.008	0.044***
domestic			2.646***	3.268***			2.681***	3.357***			2.167***	2.123***			2.315***	2.189***
language			0.089	0.093			0.239**	0.165			0.055	0.153***			-0.053	0.182***
outer o							0.708	0.635			1.286***	1.179***			0.823***	1.033***
outer d			1.607***	1.428***							1.190***	1.236***			1.518***	1.479***
island o			1.557***	1.580***			-0.647***	-0.651***			0.328***	0.314***			0.284***	0.290***
island d			-0.211	-0.110			0.673**	1.387***			0.233***	0.262***			0.172**	0.201***
euro rel			-0.096	-0.257**			-0.108	-0.028			-0.051	-0.504***			0.010	-0.483***
precip rel			0.191***	0.191***			0.677***	0.612***			-0.208***	-0.169***			-0.043	-0.111***
temp rel			0.039**	0.041***			-0.002	-0.009			-0.030***	-0.023***			-0.012*	-0.031***
unempl o			-0.020	-0.095			-0.007	-0.027			0.409***	0.423***			0.071**	0.264***
unempl d			-0.442***	-0.452***			-0.321***	-0.412***			-0.165***	-0.091**			-0.173***	-0.191***
schen rel			1.017***	0.968***			1.133***	0.748***			0.229***	0.408***			0.426***	0.706***
disp inc o		-1.915***				-0.239				-0.753***				-1.008***		
disp inc d		1.350***				-0.141				-0.123				0.391***		
h edu o			0.028	0.056			-0.741***	-0.502**			0.534***	0.576***			0.649***	0.606***
rd exp o			-0.172***	-0.178***			-0.000	-0.062			0.137***	0.121***			-0.054*	-0.052**
new eu o																
agr sh o																
metro o																

metro d																
urban o																
urban d																
lab mar long				-0.065***				0.066				-0.207**				-0.692***
lab mark med				0.816***				0.264***				0.336***				0.982***
lab mark short				0.875***				1.535***				1.659***				0.944***
emp y o																
emp sen o																
emp w o																
emp mobil o																
Constant	-12.97***	-7.90***	-9.91***	-9.80***	-9.510***	-5.396**	-1.895	-2.705	-10.14***	0.27	-13.03***	-12.08***	-10.42***	-2.78*	-17.60***	-14.64***
Observations	147,341	130,018	134,514	134,514	147,223	129,902	99,774	99,774	501,400	391,713	289,803	289,803	501,536	391,833	328,430	328,430
R-squared	0.363	0.445	0.379	0.374	0.307	0.328	0.413	0.423	0.445	0.425	0.611	0.613	0.440	0.419	0.567	0.589

Robust standard errors for significance tests; *** p<0.01, ** p<0.05, * p<0.1; the PPML estimation results.

Source: Own elaboration.

Table 3.5: Estimation results of the econometric model of migration for metropolitan and non-metropolitan regions in 2010-2018

	dependent variable: $Migration\ flow_{ijt}$															
	Outflows –metropolitan regions				Inflows –metropolitan regions				Outflows – non-metropolitan regions				Inflows –non-metropolitan regions			
VARIABLES	(1)	(3)	(5)	(9)	(1)	(3)	(5)	(9)	(1)	(3)	(5)	(9)	(1)	(3)	(5)	(9)
popul_o	0.753***	0.650***	0.331***	0.557***	0.844***	0.867***	0.933***	0.984***	0.810***	0.837***	0.781***	0.911***	0.798***	0.804***	0.760***	0.956***
popul_d	0.774***	0.783***	0.784***	0.863***	0.770***	0.600***	0.640***	0.720***	0.794***	0.785***	0.664***	0.776***	0.688***	0.727***	0.695***	0.787***
dist	-1.366***	-1.517***	-0.906***	-0.919***	-1.403***	-1.531***	-0.865***	-0.869***	-1.680***	-1.670***	-0.995***	-1.062***	-1.665***	-1.669***	-1.008***	-1.082***
gdp_pc_o	-0.086*		-0.570***	-0.695***	-0.502***		-0.279**	-0.413***	-0.536***		-0.639***	-0.646***	-0.107***		-0.431***	-0.536***
gdp_pc_d	0.083		0.497***	0.423***	0.253***		0.410***	0.222***	0.330***		0.363***	0.185***	0.033		0.168***	0.056
pop_den_rel			-0.006	-0.018			0.093***	0.064***			0.139***	0.109***			-0.063***	-0.081***
domestic			2.543***	2.533***			2.407***	2.485***			2.501***	2.431***			2.517***	2.425***
language			-0.195***	-0.240***			-0.133**	-0.208***			-0.066*	-0.123***			-0.122***	-0.197***
outer_o							0.966***	1.214***			1.187***	1.442***			0.257	0.541***
outer_d			1.389***	1.371***							1.424***	1.768***			2.063***	2.279***
island_o			0.026	0.150			0.350**	0.343**			0.486***	0.506***			0.165*	0.265**
island_d			0.252*	0.316**			0.212*	0.337***			0.159**	0.111			0.130	0.105
euro_rel			0.191***	0.097			0.012	0.078			-0.104**	-0.010			0.097**	0.108*
precip_rel			-0.037	-0.064			-0.070	-0.221***			0.012	-0.106**			0.193***	0.106**
temp_rel			-0.039***	-0.041***			0.030***	-0.001			0.010	-0.018**			-0.012*	-0.025***
unempl_o			0.176***	0.031			0.256***	0.261***			-0.087**	0.008			-0.032	-0.047
unempl_d			-0.059	-0.075			-0.366***	-0.411***			-0.246***	-0.300***			-0.247***	-0.255***
schen_rel			0.565***	0.536***			0.543***	0.389***			0.575***	0.324***			0.477***	0.273***
gdp_pc_rel																
disp_inc_o		-0.069				-1.189***				-1.043***				-0.215***		
disp_inc_d		0.208*				1.171***				0.714***				-0.027		
h_edu_o			0.723***	0.904***			0.191*	0.659***			0.019	0.518***			0.409***	0.782***
rd_exp_o			0.186***	0.097			-0.092*	-0.061			-0.030	-0.061*			-0.025	-0.059*
new_eu_o																
agr_sh_o																

metro o												-0.398***				
metro d												-0.077				
urban o				-0.406***				-0.199*								-0.372***
urban d				-0.085				-0.101								-0.131**
lab mar long																
lab mark med																
lab mark short																
emp y o																
emp sen o																
emp w o																
emp mobil o																
Constant	-8.33***	-7.34***	-10.23***	-14.60***	-10.28***	-6.25***	-12.37***	-16.70***	-8.51***	-4.90***	-11.19***	-16.09***	-6.39***	-4.46***	-12.70***	-17.79***
Observations	175.856	141.463	112.065	83.491	175.718	141.337	110.006	82.897	489.572	393.715	328.098	199.799	489.710	393.841	330.157	200.393
R-squared	0.469	0.420	0.530	0.580	0.447	0.448	0.494	0.605	0.488	0.420	0.590	0.673	0.465	0.407	0.543	0.619

Robust standard errors for significance tests; *** p<0.01, ** p<0.05, * p<0.1; the PPML estimation results.

Source: Own elaboration.

Table 3.6: Estimation results of the econometric model of migration flows between regions of the same country during 2010-2018 (domestic migration)

	dependent variable: $Mgration\ flow_{ijt}$									
VARIABLES	(1)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
popul o	0.585***	0.594***	0.593***	0.531***	0.576***	0.689***	0.409***	0.691***	0.531***	0.068
popul d	0.594***	0.537***	0.557***	0.583***	0.555***	0.517***	0.464***	0.785***	0.583***	0.576***
dist	-1.044***	-1.146***	-1.223***	-1.151***	-1.147***	-1.221***	-1.144***	-1.193***	-1.151***	-1.162***
gdp pc o	-0.137*		-0.054	-0.275**		-0.604**	-0.248**	-0.328***	-0.275**	-0.295**
gdp pc d	0.132**		0.141*	0.103		0.646***	0.049	-0.021	0.103	0.199**
pop den rel			0.031	0.056*	0.080***	0.024	0.050	0.056**	0.056*	0.045
language			0.941***	0.920***	0.874***	1.121***	0.933***	0.873***	0.920***	0.959***
outer o			1.400***	1.404***	1.390***	1.164***	1.428***	1.713***	1.404***	1.299***
outer d			1.675***	1.635***	1.673***	1.423***	1.714***	1.731***	1.635***	1.604***
island o			0.434***	0.561***	0.711***	0.425***	0.608***	0.547***	0.561***	0.541***
island d			0.358***	0.156	0.241**	0.357**	0.180	0.223	0.156	0.157
precip rel			0.092	-0.025	0.031	0.119	-0.025	-0.152*	-0.025	0.003
temp rel			0.015	0.009	0.006	0.018	0.009	-0.025	0.009	0.014
unempl o			0.133*	0.177*		-0.047	0.165*	0.289***	0.177*	0.208**
unempl d			-0.047	-0.103		0.133	-0.143*	-0.199***	-0.103	-0.117
disp inc o		-0.995***								
disp inc d		0.672***								
h edu o				0.159**	0.126*	0.559***	0.105	0.662***	0.159**	0.261***
rd exp o				0.079*	0.059*	-0.054	0.092**	0.076*	0.079*	0.052
new eu o					0.317***					
agr sh o						-0.050				
metro o							0.263***			
metro d							0.319***			
urban o								-0.357***		
urban d								-0.285***		
emp y o										-0.318***
emp sen o										0.075
emp w o										0.105
emp mobil o										0.476
Constant	-4.307***	0.149	-4.751***	-6.408***	-5.712***	-7.457***	-3.076***	-13.37***	-6.408***	-2.553
Observations	47,260	35,968	41,688	28,330	30,374	18,388	28,330	22,151	28,330	27,332
R-squared	0.460	0.444	0.438	0.407	0.409	0.418	0.410	0.528	0.407	0.411

Robust standard errors for significance tests; *** p<0.01, ** p<0.05, * p<0.1; the PPML estimation results.

Source: Own elaboration.

Table 3.7: Estimation results of the econometric model of migration flows between regions of different countries in 2010-2018 (international migration)

	dependent variable: $Migration\ flow_{ijt}$									
VARIABLES	(1)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
popul o	0.977***	0.916***	1.063***	1.315***	1.129***	1.315***	1.122***	1.376***	1.351***	1.076***
popul d	0.862***	0.835***	0.970***	1.001***	0.879***	1.036***	0.875***	0.878***	0.984***	0.974***
dist	-0.371***	-0.366***	-0.297***	-0.324***	-0.324***	-0.330***	-0.330***	-0.356***	-0.292***	-0.349***
gdp pc o	-0.507***		-0.473***	-0.308***		-0.147***	-0.334***	-0.378***	0.370***	-0.167***
gdp pc d	0.308***		0.127***	0.358***		0.388***	0.276***	0.242***	0.829***	0.287***
pop den rel			0.024**	0.037***	0.150***	0.061***	0.024**	-0.006	0.036***	0.072***
language			-0.063**	-0.028	0.225***	0.073**	0.000	-0.185***	0.232***	-0.002
outer o			-0.146	0.088	0.240*	-0.020	0.121	0.042	0.122	-0.410***
outer d			1.177***	1.269***	0.745***	1.609***	1.464***	1.209***	1.531***	1.237***
island o			0.166**	-0.002	-0.340***	-0.071	0.100	0.022	0.020	0.435***
island d			0.276***	0.207***	0.282***	-0.097	0.143**	0.150*	0.210***	0.298***
euro rel			0.154***	0.071**		0.143***	-0.023	0.069*	-0.533***	0.125***
precip rel			0.116***	0.012	-0.006	0.002	0.027	-0.025	0.054*	-0.000
temp rel			0.010***	0.002	-0.072***	0.002	0.008*	-0.002	-0.003	-0.000
unempl o			-0.257***	-0.195***		-0.182***	-0.186***	-0.273***	-0.149***	-0.136***
unempl d			-0.579***	-0.609***		-0.664***	-0.670***	-0.578***	-0.571***	-0.594***
schen rel			0.199***	0.241***		0.475***	0.355***	0.034	0.411***	0.194***
disp inc o		-0.719***								
disp inc d		0.715***								
h edu o				0.465***	0.484***	0.454***	0.271***	0.670***	0.504***	0.016
rd exp o				-0.193***	-0.081***	-0.318***	-0.192***	-0.265***	-0.230***	-0.233***
new eu o					0.910***					
agr sh o						-0.265***				
metro o							0.470***			
metro d							0.385***			
urban o								0.040		
urban d								0.421***		
lab mar long									-0.413***	
lab mark med									1.086***	
lab mark short									1.758***	
emp y o										-0.115***
emp sen o										-0.120***
emp w o										4.310***
emp mobil o										0.353**
Constant	-21.04***	-19.00***	-23.61***	-26.37***	-24.62***	-25.12***	-21.71***	-26.00***	-22.79***	-39.07***
Observations	618,168	499,210	557,182	411,833	478,946	342,728	411,833	261,139	411,833	387,164
R-squared	0.126	0.115	0.169	0.218	0.138	0.251	0.232	0.197	0.303	0.228

Robust standard errors for significance tests; *** p<0.01, ** p<0.05, * p<0.1; the PPML estimation results.

Source: Own elaboration.

4 Key questions for subsequent tasks

4.1 Task 2. Pan-European systemic analysis

The matrices that have been put together provide ample possibilities for use in the synthetic considerations of Task 2's framework. Our methodology ensured that we could fill in all the cells of the matrix for all considered instants of time.

At the same time, our analysis of migration flows demonstrated the existence of several essential elements to be considered in the elaboration of standardized matrices, and especially in the interpretation of Task 2's results:

- International and domestic migrations differed very significantly in the nature of the original data, the territorial distribution of many indicators, and the explanatory factors. The summary results may in effect result from the various situations in the framework of the two basic migration types. They may present a kind of virtual "intermediate situation". While the reality remains identifiable in the interpretation of the matrices and interaction maps based on real data, an essential difficulty might arise for the analysis of standardized matrices.
- Despite the application of appropriate indicators, the magnitude of the NUTS 2 units in certain countries still exerted a significant influence on data for the distribution of migration flows. This influence was probably bigger here than for some other kinds of flows. This was of special importance for internal migrations, for which the data obtained account for a bigger or smaller part of the local movements, depending upon the surface of the statistical units. Again, this might pose a more difficult problem for standardized metrics and synthetic indicators.
- The region typologies elaborated for migration flows enabled us to indicate a number of territorial clusters and groups of dispersed regions, which deserve a more profound analysis in synthetic assessments (including comparison with other kinds of flows).

4.2 Task 3. Scenarios

Our study of migration flows, carried out in association with an analysis of explanatory factors, confirmed the correctness of the case studies (outflow and inflow maps are presented below for each case study) selected for more in-depth investigations in the framework of Task 3. These cases represent different migration types, which ought to help us identify the local interdependences and formulate recommendations for territorial policies.

In the context of the scenarios selected for analysis, the study of migration flows may point out problems worthy of further assessment, especially in the course of qualitative analysis according to the selected scenarios.

“Brexit” scenario

- Migration flows from Central-Eastern Europe to the UK had been decreasing before the formal Brexit.
- The UK was not homogeneous in its migration connections with the regions of Europe. There were distinct differences (in terms of various indicators and also in the typological image) between: a) central England together with London, b) the remaining areas of England and Wales, c) Scotland, and d) Northern Ireland.

Figure 4.1: Migration outflow from Eastern and Midland region, Ireland

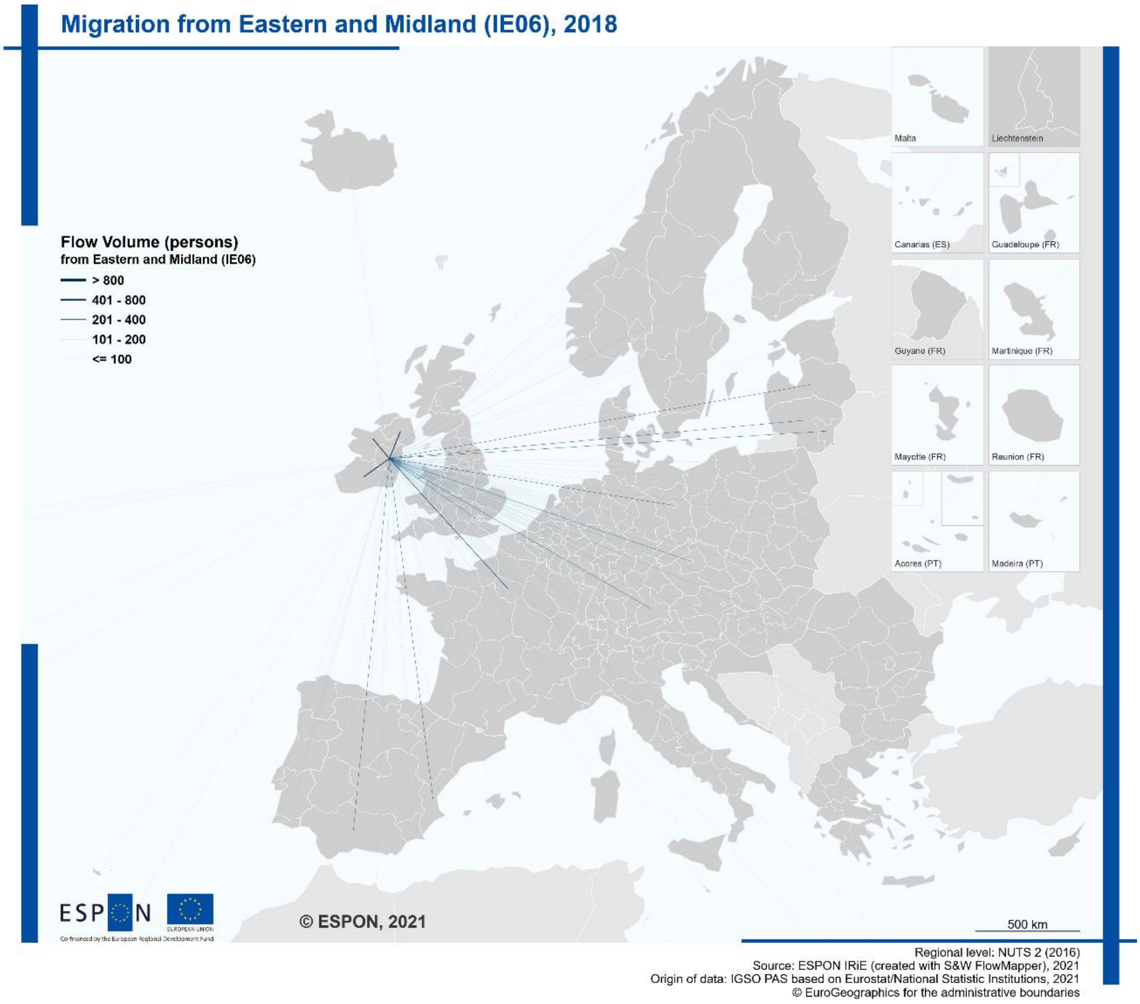
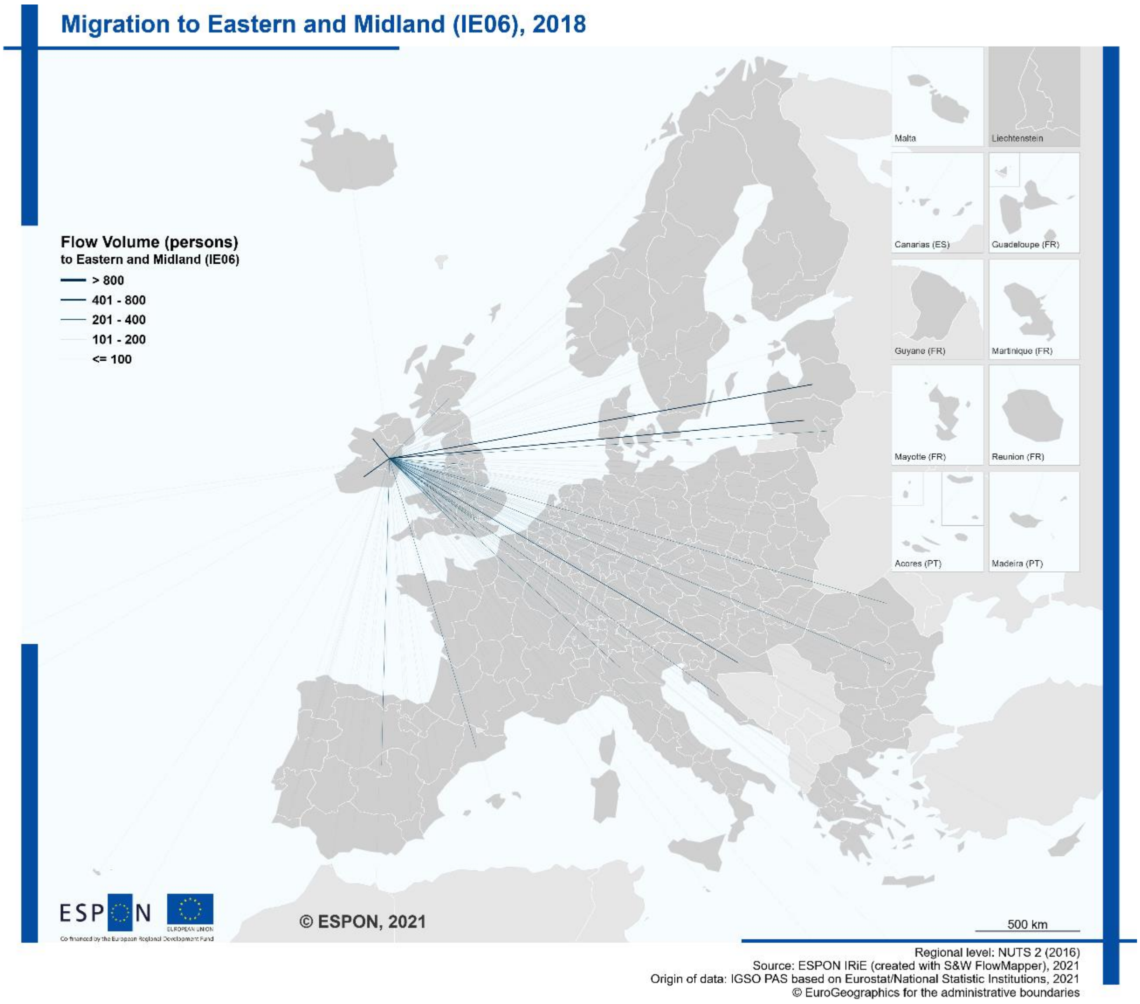


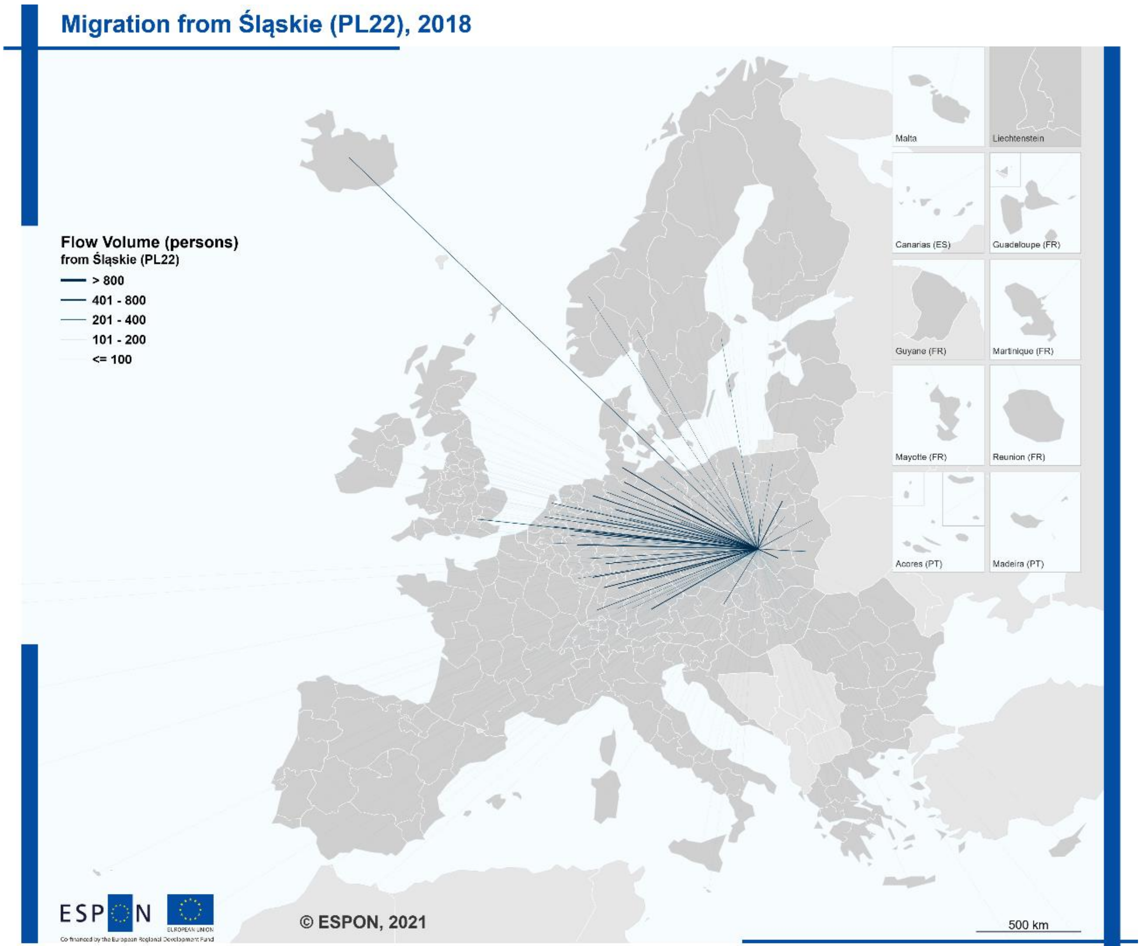
Figure 4.2: Migration inflow to Eastern and Midland region, Ireland



“Green Deal” scenario

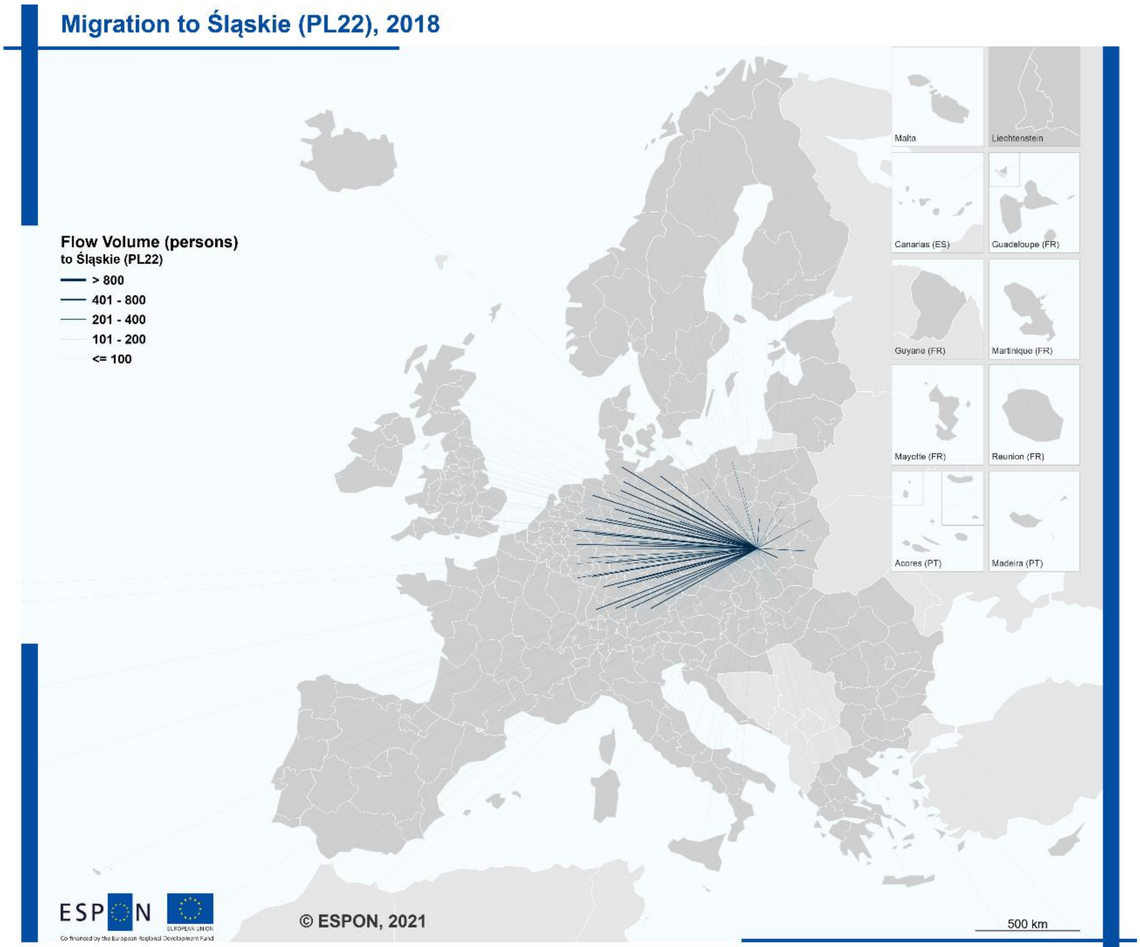
- Certain regions with high shares of traditional industries (especially coal extraction and coal-based energy production) were already characterized by a distinct migration outflow. In this context the changes brought about by energy-related transformation might be treated as an accelerator of changes in the setting of extant migration flows.

Figure 4.3: Migration outflow from Śląskie region, Poland



Regional level: NUTS 2 (2016)
Source: ESPON IRiE (created with S&W FlowMapper), 2021
Origin of data: IGSO PAS based on Eurostat/National Statistic Institutions, 2021
© EuroGeographics for the administrative boundaries

Figure 4.4: Migration inflow to Śląskie region, Poland



Regional level: NUTS 2 (2016)
Source: ESPON IRiE (created with S&W FlowMapper), 2021
Origin of data: IGSO PAS based on Eurostat/National Statistic Institutions, 2021
© EuroGeographics for the administrative boundaries

"New Globalization" scenario

- The analysis of migration flows is not directly related to the scenario founded on trade relations.

Figure 4.5: Migration outflow from Zuid-Holland region, the Netherlands

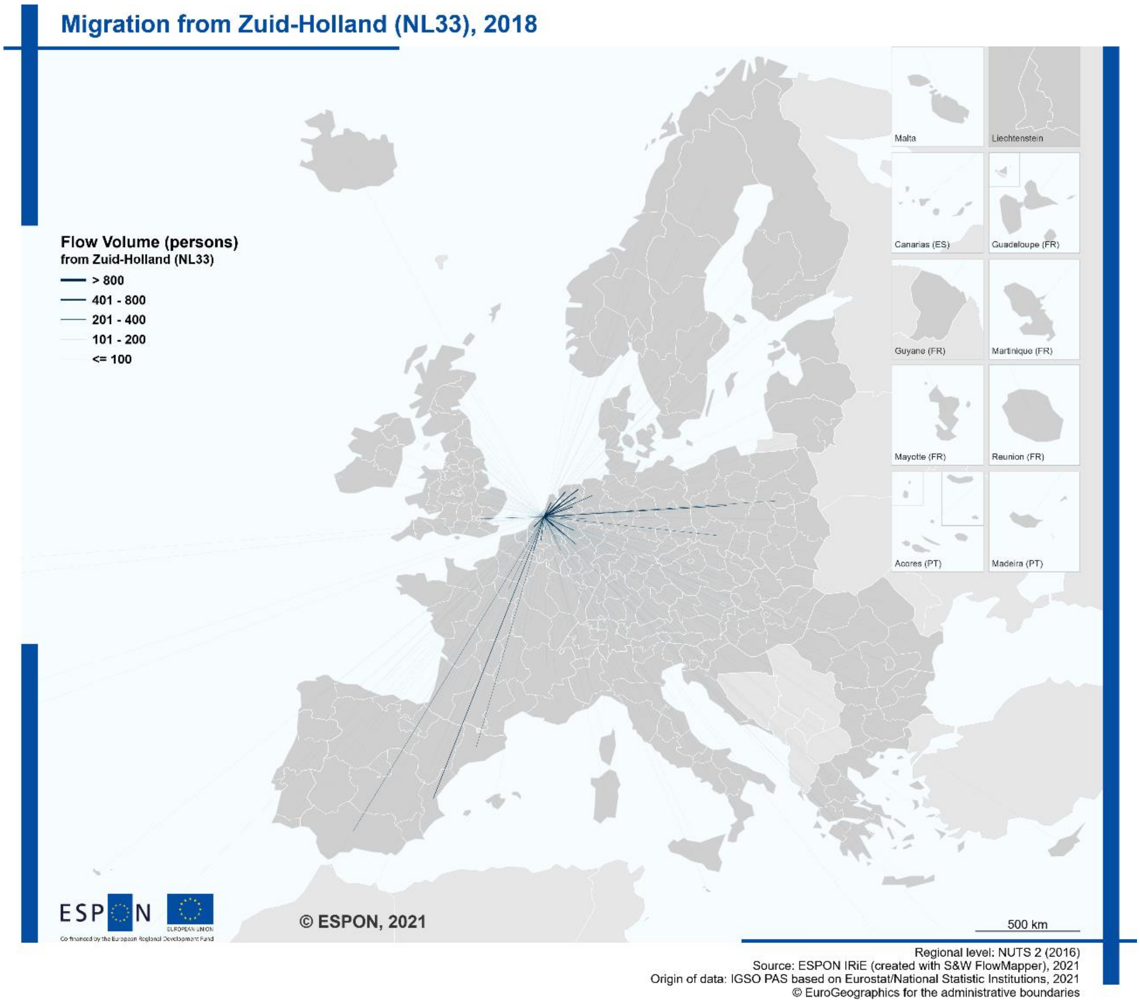
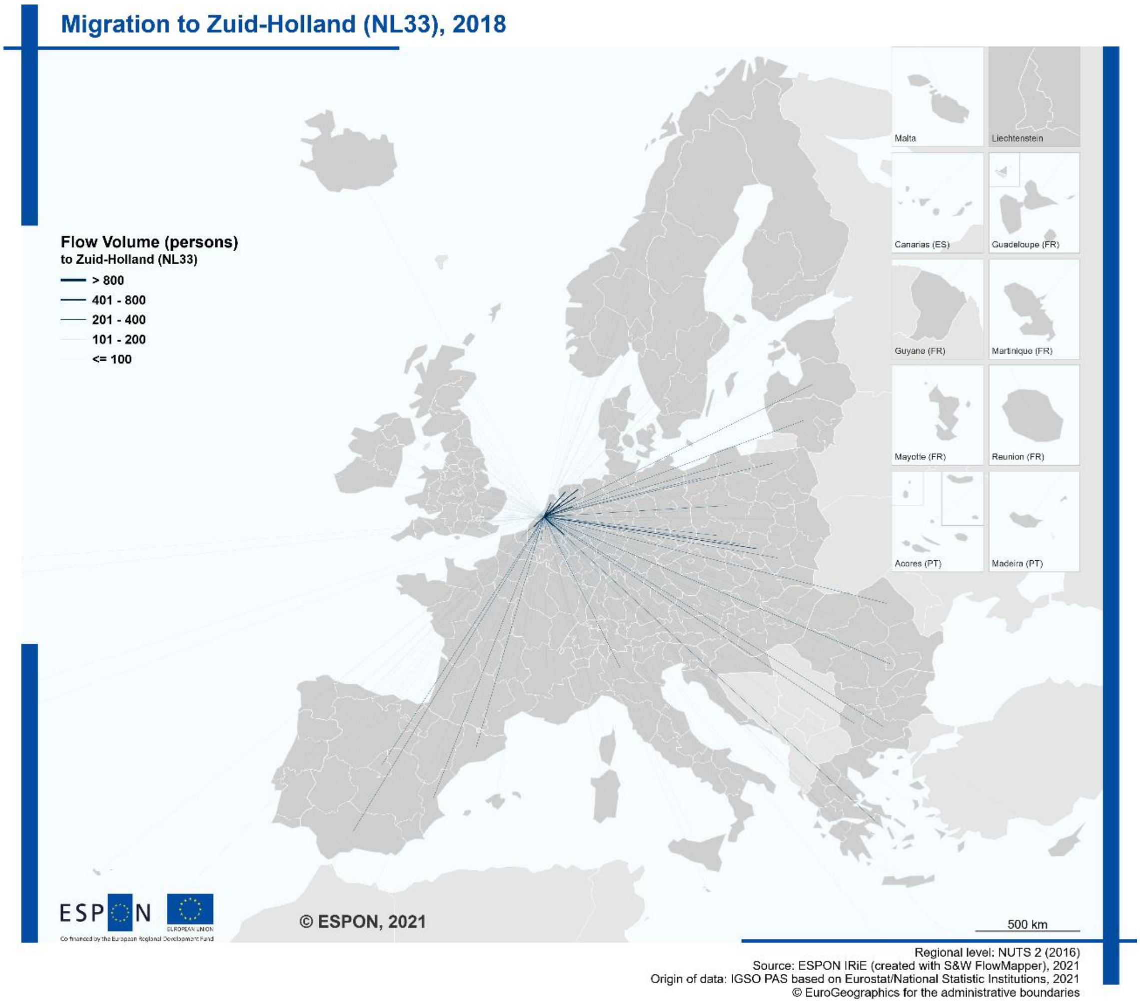


Figure 4.6: Migration inflow to Zuid-Holland region, the Netherlands



“Covid-19” scenario

- Because of the dominance of internal migrations in the total volume of the European migrations, changes in labour markets (remote working) may influence the image on the continental scale. Potential Covid-19-related extensions of job-commuting distances may lead to local migrations (also to return migrations) to zones more distant from metropolises. Such migrations will register first as a statistical increase at the NUTS 2 level (especially in countries, like Germany and the UK, where such units are small).

Figure 4.7: Migration outflow from Comunidad Foral de Navarra region, Spain

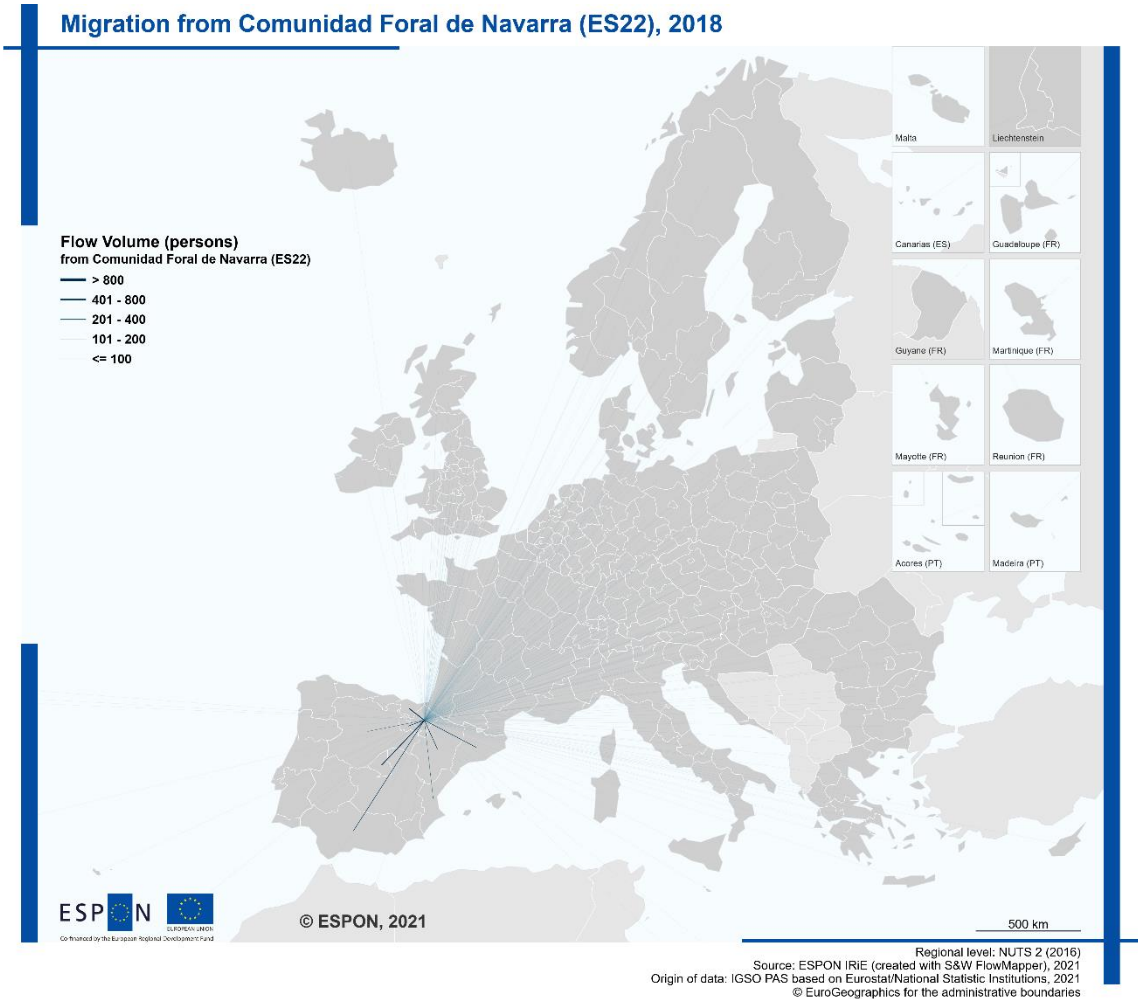
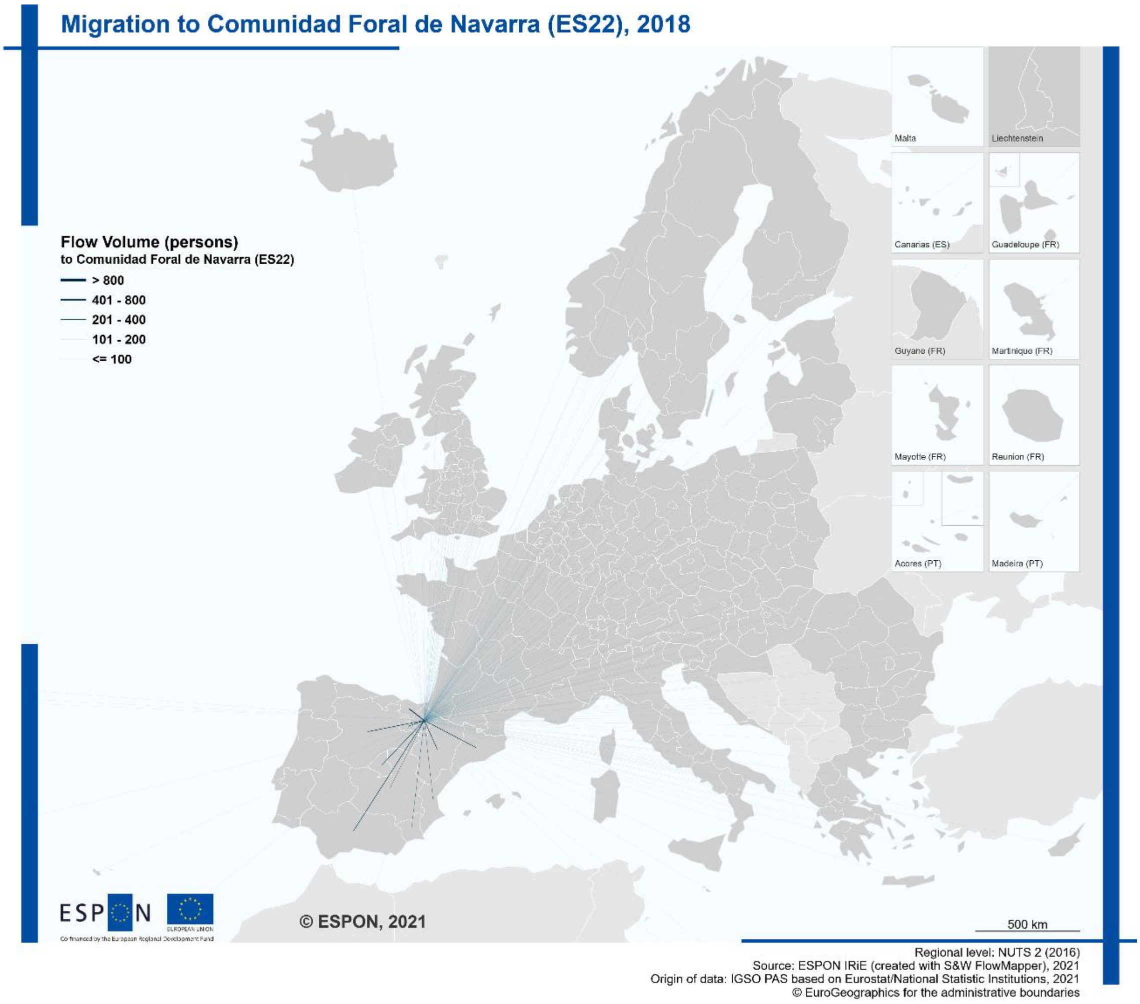


Figure 4.8: Migration inflow to Comunidad Foral de Navarra region, Spain



4.3 Task 4. Policy implications

The analysis of migration flows may constitute a good basis for the formulation of policy recommendations at the European, national, and regional levels. Of special importance in this context are:

- Conclusions connected with the quality of the data available (see below). Improvement of quality is fundamental for the correct assessment of regional potentials (recognition of the true populations, labour force resources, etc.).
- Conclusions relative to the existence of sui generis migration hubs in some cities of the peripheral zones (especially in Central-Eastern Europe). These cities might require a particular (flexible) policy for the provision of public service. Likewise, their transport-related policy might be different, to shape their position appropriately within these countries and in Europe.
- Conclusions with respect to the transitory character of the high-intensity migrations over definite directions that occur once labour markets are opened up. This is important for

predictions concerning the labour market and the number of persons using public services.

5 Recommendations for data providers to improve data quality

European migration statistics by country of previous residence and country of next residence are incomplete and come from different sources and surveys, resulting in differences in the data on flows between pairs of countries. There is no O-D data at the regional level. Eurostat collects migration data from NSIs using various data sources, according to national availability and practice. Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection provides the basis for data capture and dissemination.

At a time of huge increases in the mobility of people in Europe with freedom of movement (migration, circular migrant flows, cross-country commuting), migration statistics should be a priority for the development of public statistics. Their coherent collection and compilation should be compulsory for EU countries.

Efforts should be made at the EU level (Eurostat) to widen the scope of the data collected, which will increase the statistical burden but also allow the development of regionally targeted policies. The extension of statistics should:

- improve the quality of data on C2C (long-term) population flows,

- include short-term migration flows and expand the scope of the data made available so that it includes C2C migration,

- include the age and gender of migrants in C2C flows,

- acquire and make available regional data on flows within countries,

- discuss the development of a system for recording population flows between European countries (e.g. like the existing system for foreign trade), which would allow current demographic trends in the EU to be monitored and demographic policies to be pursued accordingly.

Ultimately, European migration statistics should be available to citizens, decisionmakers, and researchers without the enormous effort required to compare the various data sources and forms of data publication by the NSI.

6 Conclusions

The dominant direction of migration-related connections in the ESPON space was determined by East-West flows, i.e. those from countries that joined the EU in 2004 (the biggest ones from Poland and Romania). These flows were mainly oriented at Germany, the UK, Spain, and Italy. The role of Germany had distinctly increased in terms of the share of migration flows (increase by 5 percentage points in the years 2010-2018). The negative balance of migrations, characteristic of countries of the so-called “new EU” (the biggest source of migrants for the countries of Western Europe), concerned also the countries of the so-called “old EU”, namely Spain, Portugal, Greece, and Italy. Ample migration outflow from these countries can be associated with the consequences of the financial crisis of 2008.

Likewise, at the regional level, the dominant direction of migration flows was the East-West. Analysis of these flows demonstrated that interregional flows between the countries of Central Europe were of marginal significance with regard to the totality of flows within the ESPON space. The majority of the strongest migration flows in Europe was constituted by internal movements in the countries considered. They were concentrated within urban agglomerations and their surroundings (Functional Urban Areas: e.g. Ruhr Basin, Greater London). This may have resulted from, for instance, suburbanization. These interregional flows occurred on the largest scale in Germany, and were due not only to German society’s mobility but also to the size of the country’s NUTS 2 units.

Cities together with their functional neighbourhood were characterized, as well, by the highest positive net migration balance, even in the countries with a negative, or very negative, migration balance: e.g., Warsaw (featuring one of the highest values of the indicator in Europe), Berlin, Budapest, Bucharest, Sofia, and Athens. A compact area with a distinctly positive migration balance stretched from the Netherlands through western Germany, Austria, Switzerland, and down to central Italy. A clearly positive migration balance was also observed in the island regions of Spain (Baleares, Canary Islands) and France (Corsica). Yet no analogue was observed on the Italian (Sicily, Sardinia), Portuguese (Azores, Madeira), or Greek islands. A negative balance of migrations concerned primarily the regions of Poland, Romania, Bulgaria, and Lithuania. Moreover, the values of the net migration balance in particular countries displayed a very high degree of internal differentiation (e.g. in the UK).

It was internal flows (i.e. taking place within one country) that played a leading role in the totality of migration flows. Only some of the countries considered had smaller domestic than international migrations (Lithuania, Croatia, Bulgaria, Poland, Portugal, Slovakia, Ireland, Romania). In the remaining cases internal migrations were bigger than external ones, in some cases even several times over (the UK, Finland, Sweden).

In many regions of the ESPON space the level of migratory imbalance (negative or positive) was quite limited (e.g. in France, Spain, Portugal, the Scandinavian countries, Ireland, Greece, Slovakia). From the point of view of intra-European migrations the population there was stable. In the years 2010-2018 the migration balance became clearly higher in only a few of European MEGAs, namely Munich, Frankfurt, London, Berlin, and Warsaw.

Distance was among the fundamental elements determining spatial relations, exerting the greatest influence on the length of migrations in peripheral regions, especially when external migrations in these regions were pronounced (Bulgaria, Romania, eastern Poland, Portugal) and when the country's geographic dimensions imposed long internal migrations (Italy, Norway). Moreover, the indicator for average migration distance showed core-periphery patterns in certain countries (the Scandinavian countries, Switzerland, the UK, Hungary, Slovakia).

Analysis of the mutual relation of the indicator for dependence upon migration inflow and outflow distance showed that the eastern part of the ESPON space was dominated by relative significance of distance in inflow as compared to outflow. Emigration from these areas was conditioned on other factors. At the same time, immigration occurred on a lower scale and was more regional. On the other hand, most of the countries of Western Europe (including Spain, the UK, and Sweden) presented the inverse situation. Distance appeared there mainly as affecting migration outflow. Inflows took place from many, sometimes quite dispersed origins.

The concentration of migration relations with respect to population brought out national patterns of high concentration in internal migrations and a small number of metropolises attracting migrants (e.g. Hungary, Austria, Czechia, the Netherlands, Norway, Sweden, Denmark, Finland). Lower values of the indicator were observed in bigger countries, in both Western and Central-Eastern Europe. The reasons may be sought in their more polycentric settlement systems, as well as in the dispersal of their international migrations. In peripheral countries the higher values of this indicator appeared usually around capital units and other metropolises (draining of migrants from the direct hinterland of the metropolises).

High selectivity of migrations (i.e. the highest shares of the strongest relations from the side of the migrants' region of origin) characterized some of the peripheral regions of the ESPON space (e.g. Norway, Sweden, Finland), along with the neighbourhoods of metropolises (e.g. Berlin, Prague, Budapest, Bucharest, Athens, Madrid). On the other hand, the highest level of external influence on migration flows (i.e. the highest share of the strongest outflow from the side of the receiving region — the inverse of the selectivity indicators) was characteristic of the cores of large metropolises. Their labour markets displayed high dependence on the inflow of migrants (e.g. London, Paris, Lisbon, Copenhagen, Berlin, Madrid, Warsaw, Budapest, Vienna, Oslo).

Strong or moderate domination of outflow coupled with high-intensity migrations was observed primarily in Romania, Bulgaria, and Greece. These areas were followed by eastern Hungary, most of eastern Germany's units, northern Finland and Norway, and central Spain. Many other areas of distinctly emigrational character (negative migration balance) featured a relatively weak general intensity of migration (Poland, southern Italy, western Spain, Portugal). Regions characterised by strong, imbalanced migration inflows constituted a compact area primarily in Germany, middle Sweden, Norway, Switzerland, Ireland, and northern Scotland. Certain metropolises — in Germany (Hamburg, Berlin), the UK (London, Liverpool), and Central-Eastern Europe (Bucharest, suburban zone of Budapest) — fell into the same group.

High intensity with simultaneous concentration of migration was observed primarily in the UK (England), western Bavaria, and western Hungary. Certain regions with high intensity of flows featured in addition low concentration (regions of southern Germany, middle Norway, southern Sweden, and some metropolitan centres, such as Berlin, Oslo, Bucharest).

In migratory intensity and average migratory distance the dominant regions featured moderate distances with simultaneous relatively high intensity (western Germany; eastern as well as western England; single Swedish, Norwegian, and Romanian units). High intensity coupled with short migration distance was observed primarily in central England, the Netherlands, southern Belgium, eastern Denmark, central Hungary, and Czechia (particularly the area surrounding Prague).

Our overall assessment of the three typologies elaborated (in connection with the analysis of the basic indicators) indicates the separate character of some of the territorial clusters in terms of flow intensity, balance, concentration, and distance of migrations.

Our general model confirmed the significance of the associations between magnitude of migration flows and income levels; the wealth and population of the receiving and origin regions; and distance. The affinity of languages (the same language or the same group of languages) was an essential driver of migration flows. The opening up of labour markets in particular regions was another. Also, the shorter the time since the opening up of the labour market, the stronger the influence (novelty effect). A bad labour market in the region of (potential) destination would constitute a barrier to flows. Regions of origin where a high percentage of the population had received higher education ought to produce more intensive migration flows. Membership in the Schengen zone constituted an essential strengthening factor for migration flows (driver), but the effect of membership in the euro zone was not that obvious.

Unlike the model for total flows, the model for inflows to new member countries of the EU and non-metropolitan regions omits the significance of economic factors (GDP per capita and the disposable income of households) for the destination region. Here, an important stimulant to the flow of migrants was internal movements, frequently undertaken for non-economic reasons. Quite similarly, the economic criteria of GDP and income at the disposal of households lost their significance in the model for outflows from regions of “old” EU member countries.

The development of the R2R matrix for ESPON space countries is challenging, particularly methodologically, if only for lack of a European source of statistics on regional migration flows. There are also multiple gaps in the raw data, each with a different span and a different solution:

- gaps in time series (beginning, end, middle, interleaved / single year, many years);
- gaps in country coverage;

- gaps in region coverage for single countries;
- lack of data in one direction (outflow or inflow);
- incomplete matrix of flows (e.g., only main migration directions : top 5, top 10).

Moreover, European statistics lack data on internal migration. This entails additional challenges: e.g. different approaches to the application of national methodologies; different territorial divisions, definitions, and data sources.

References

Albrechts L and Coppens T (2003) Megacorridors: striking a balance between the space of flows and the space of places. *Journal of Transport Geography* 11(3): 215–224.

Andersen TB and Dalgaard C-J (2011) Flows of people, flows of ideas, and the inequality of nations. *Journal of Economic Growth* 16(1): 1–32.

Backhaus A, Martinez-Zarzoso I and Muris C (2015) Do climate variations explain bilateral migration? A gravity model analysis. *IZA Journal of Migration* 4(1).

Baltagi BH and Khanti-Akom S (1990) On efficient estimation with panel data: An empirical comparison of instrumental variables estimators. *Journal of Applied Econometrics* 5(4): 401–406.

Baltagi BH, Egger P and Pfaffermayr M (2003) A generalized design for bilateral trade flow models. *Economics Letters* 80(3): 391–397.

Basher SA and Fachin S (2008) The long-term decline of internal migration in Canada: the case of Ontario. *Letters in Spatial and Resource Sciences* 1(2-3): 171–181.

Beine M and Parsons C (2015) Climatic Factors as Determinants of International Migration. *The Scandinavian Journal of Economics* 117(2): 723–767.

Beine M, Bourgeon P and Bricongne J (2019) Aggregate Fluctuations and International Migration. *The Scandinavian Journal of Economics* 121(1): 117–152.

Bell M, Blake M, Boyle P, et al. (2002) Cross-national Comparison of Internal Migration: Issues and Measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165(3): 435–464.

Blatter J (2004) From “spaces of place” to “spaces of flows”? Territorial and functional governance in cross-border regions in Europe and North America. *International Journal of Urban and Regional Research* 28(3): 530–548.

Boschma R and Iammarino S (2009) Related Variety, Trade Linkages, and Regional Growth in Italy. *Economic Geography* 85(3): 289–311.

Bredtmann J, Nowotny K and Otten S (2020) Linguistic distance, networks and migrants’ regional location choice. *Labour Economics* 65: 101–863.

Bush SR and Oosterveer P (2007) The Missing Link: Intersecting Governance and Trade in the Space of Place and the Space of Flows. *Sociologia Ruralis* 47(4): 384–399.

Castells M (1998) *End of Millenium, the Information Age: Economy, Society, Culture, Vol. III*. Cambridge, MA, Oxford, UK: Blackwell.

Castells M (2007) Communication, Power and Counter-Power in the Network Society. *International Journal of Communication* 1: 238–266.

Castles S and Miller MJ (2011) *Migracje We Współczesnym Świecie*. Warsaw: Wyd. Naukowe PWN.

Chojnicki Z (1961) Analiza przepływów towarowych w Polsce w układzie międzywojewódzkim. *Studia KPZK PAN* 1: 193.

Davis D and Gift T (2014) The Positive Effects of the Schengen Agreement on European Trade. *The World Economy* 37(11): 1541–1557.

Derudder B and Witlox F (2005) On the use of inadequate airline data in mappings of a global urban system. *Journal of Air Transport Management* 11(4): 231–237.

Egger P and Pfaffermayr M (2003) The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects. *Empirical Economics* 28(3): 571–580.

ESPON (2004) *ESPON 1.1.1: Potentials for Polycentric Development in Europe. Final Project Report*. Luxembourg: ESPON 2006 Programme. Available at: https://www.espon.eu/programme/projects/espon-2006/thematic-projects/urban-areas-nodes-polycentric-development (accessed 21 October 2021).

ESPON (2007a) *ESPON 1.4.3: Study on urban functions. Final Report*. Luxembourg: ESPON 2006 Programme. Available at: https://www.espon.eu/programme/projects/espon-2006/studies-and-scientific-support-projects/study-urban-functions (accessed 21 October 2021).

ESPON (2007b) *ESPON 3.4.1: Europe in the World. Final Report*. Luxembourg: ESPON 2006 Programme. Available at: https://www.espon.eu/programme/projects/espon-2006/coordinating-cross-thematic-projects/europe-world (accessed 21 October 2021).

ESPON (2010) *FOCI: Future Orientations For Cities. Final Report*. Luxembourg: ESPON 2013 Programme. Available at: https://www.espon.eu/programme/projects/espon-2013/applied-research/foci-future-orientation-cities (accessed 21 October 2021).

Faist T (2000) *The Volume and Dynamics of International Migration and Transnational Social Spaces*. Oxford: Clarendon Press.

Hausman JA and Taylor WE (1981) Panel Data and Unobservable Individual Effects. *Econometrica* 49(6): 1377–1398.

Isard W (1951) *Interregional and Regional Input-Output Analysis: A Model of a Space-Economy.* Cambridge: The MIT Press.

Isard W (1965) *Metody Analizy Regionalnej*. Warszawa: KPZK PAN, PWN.

Kanbur R and Rapoport H (2005) Migration selectivity and the evolution of spatial inequality. *Journal of Economic Geography* 5(1): 43–57.

Kritz MM and Zlotnik H (1992) *Global Interactions. Migration Systems, Process and Policies*. New York: Oxford University Press.

Kupiszewski M and Kupiszewska D (2010) *Reference scenarios. Report on Deliverable 5. DEMIFER. Demographic and Migratory Flows affecting European Regions and Cities. Applied Research Project 2013/1/3.* ESPON.

Leibenstein H and Tinbergen J (1966) Shaping the World Economy: Suggestions for an International Economic Policy. *The Economic Journal* 76(301): 92–95.

Leontief WW (1956) Factor Proportions and the Structure of American Trade. *The Review of Economics and Statistics* 38(4): 386.

López-Bazo E and Motellón E (2018) Firm exports, innovation and the regional dimension in Spain. *Regional Studies* 52(4): 490–502.

Manzoor W, Safdar N and Mahmood HZ (2021) A gravity model analysis of international migration from BRIC to OECD countries using Poisson Pseudo-maximum likelihood Approach. *Heliyon* 7(6): e07357.

Marek Okólski (2004) *Demografia Zmiany Społecznej*. Warsaw: Scholar.

Mayda AM (2010) International migration: a panel data analysis of the determinants of bilateral flows. *Journal of Population Economics* 23(4): 1249–1274.

Molloy R, Smith CL and Wozniak A (2011) Internal Migration in the United States. *Journal of Economic Perspectives* 25(3): 173–196.

Parysek J (1997) *Podstawy Gospodarki Lokalnej*. Poznań: UAM.

Pries L (1999) New Migration in Transnational Spaces. In: Pries L (ed.) *Migration and Transnational Social Spaces*. Aldershot: Ashgate, pp. 1–35.

Qi W, Deng Y and Fu B (2019) Rural attraction: The spatial pattern and driving factors of China's rural in-migration. *Journal of Rural Studies*.

Reades J and Smith DA (2014) Mapping the "Space of Flows": The Geography of Global Business Telecommunications and Employment Specialization in the London Mega-City-Region. *Regional Studies* 48(1): 105–126.

Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection and repealing Council Regulation (EEC) No 311/76 on the compilation of statistics on foreign workers (2007).

Rietveld P (2001) Obstacles to openness of border regions in Europe. In: van Geenhuizen M and Ratti R (eds) Gaining Advantage from Open Borders: An Active Space Approach to Regional Development. Ashgate: Aldershot.

Salt J (2008) Managing new migrations in Europe: Concept and reality in the ICT sector. In: Bonfazi C, Okólski M, Schoorl J, et al. (eds) *International Migration in Europe. New Trends and New Methods of Analysis*. Amsterdam: Amsterdam University Press.

Santos-Silva JMC and Tenreyro S (2006) The Log of Gravity. *Review of Economics and Statistics* 88(4): 641–658. DOI: 10.1162/rest.88.4.641.

Serlenga L and Shin Y (2021) Gravity models of interprovincial migration flows in Canada with hierarchical multifactor structure. *Empirical Economics* 60(1): 365–390.

Stewart JQ (1948) Demographic Gravitation: Evidence and Applications. *Sociometry* 11(1/2): 31.

Stillwell J (2005) Inter-regional migration modelling: a review and assessment. In: *Paper presented at the 45th congress of the European regional science association (ERSA),* 2005.

Stock JH, Wright JH and Yogo M (2002) A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* 20(4): 518–529.

Stumpner S (2019) Trade and the geographic spread of the great recession. *Journal of International Economics* 1(19): 169–180.

Tapinos G (1994) Handel zagraniczny, technologia i migracja. In: Golinowska S and Marek E (eds) *Studia O Procesach Migracji Zagranicznych, Ekonomiczne Aspekty Procesów Migracyjnych*. Warsaw: Instytut Pracy i Spraw Socjalnych, pp. 31–40.

Thissen M, van Oort F and Diodato D (2013) Integration and Convergence in Regional Europe: European Regional Trade Flows from 2000 to 2010. In: *53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy,"* Palermo, 2013.

Timmer MP, Dietzenbacher E, Los B, et al. (2015) An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production. *Review of International Economics* 23(3): 575–605.

Ullman EL (1957) *American Commodity Flow*. University of Washington Press: Seattle.

van Bergeijk P and Brakman S (2010) *The Gravity Model in International Trade*. Cambridge: Cambridge University Press.

van der Linden JA and Oosterhaven J (1995) European Community Intercountry Input–Output Relations: Construction Method and Main Results for 1965–85. *Economic Systems Research* 7(3): 249–269.

Verwiebe R, Wiesböck L and Teitzer R (2014) New forms of intra-European migration, labour market dynamics and social inequality in Europe. *Migration Letters* 11(2): 125–136.

Wang W and Li Z (2019) The evolution of China's interregional coal trade network, 1997–2016. *Physica A: Statistical Mechanics and its Applications* 536(2029): 120–974.

Wang Y, Dong L, Liu Y, et al. (2019) Migration patterns in China extracted from mobile positioning data. *Habitat International* 86: 71–80.

Wesselbaum D and Aburn A (2019) Gone with the wind: International migration. *Global and Planetary Change* 178: 96–109.

Zakir Saadullah Khan M and Ismail Hossain M (2010) A Model of Bilateral Trade Balance: Extensions and Empirical Tests. *Economic Analysis and Policy* 40(3): 377–391.

Zaucha J, Komornicki T, Böhme K, et al. (2014) Territorial Keys for Bringing Closer the Territorial Agenda of the EU and Europe 2020. *European Planning Studies* 22(2): 246–267.

Zhang KH and Song S (2003) Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses. *China Economic Review* 14(4): 386–400.

Zhang Y, Long H, Ma L, et al. (2018) Analysis of rural economic restructuring driven by e-commerce based on the space of flows: The case of Xiaying village in central China. *Journal of Rural Studies*.

Zhu R, Wang Y, Lin D, et al. (2021) Exploring the rich-club characteristic in internal migration: Evidence from Chinese Chunyun migration. *Cities* 114: 103–198.

Zipf GK (1946) The P1 P 2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review* 11(6): 677.

Co-financed by the European Regional Development Fund

Inspire Policy Making with Territorial Evidence

espon.eu

ESPON 2021

ESPON EGTC
4 rue Erasme, L-1468 Luxembourg
Grand Duchy of Luxembourg
Phone: +352 20 600 280
Email: info@espon.eu
www.espon.eu

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.