



RE M I N D E R

ROLE OF EUROPEAN MOBILITY AND ITS IMPACTS
IN NARRATIVES, DEBATES AND EU REFORMS

MONITORING MIGRATION WITHIN THE EU WITH EXISTING DATA

MAPPING PAPER

Authors: Veronika Fajth
Melissa Siegel
Vittorio Bruni
Tanta Gelashvili

Published: April 2018



REMINDER

ROLE OF EUROPEAN MOBILITY AND ITS IMPACTS
IN NARRATIVES, DEBATES AND EU REFORMS

Monitoring migration within the EU with existing data

Mapping paper

Authors: Veronika Fajth, Melissa Siegel, Vittorio Bruni, Tamta Gelashvili

Submitted: 15 December 2017

Paper prepared as part of the REMINDER project

www.reminder-project.eu

Correspondence address:

Veronika Fajth
Maastricht Graduate School of Governance
UNU-MERIT | Maastricht University
Boschstraat 24, 6211 AX Maastricht, The Netherlands
veronika.fajth@maastrichtuniversity.nl

Acknowledgements: the authors would like to thank Roland Hosner for his helpful comments and suggestions throughout this project.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727072

Abstract

The aim of this paper was to provide a discussion and overview of the main databases available to understand mobility within the European Union. Following some key questions of interest, the paper maps existing data sources and evaluates them for their usefulness and quality in supporting intra-EU migration research. Eurostat's database on population statistics, the Labour Force Survey (both core and 2008/2014 ad-hoc modules), the migration databases of UNDESA and OECD, as well as the EIMSS survey and the special wave 72.5 of the Eurobarometer are the among the most useful sources identified at the regional level. Despite a general trend of improving European mobility statistics, some challenges and limitations regarding the measurement of intra-EU mobility seem to persist. Firstly, missing data on migrants' previous country of residence and/or lack of cross-tabulation opportunities make it impossible to assess those movements that truly takes part within the external borders of the EU and/or the share of EU versus third-country nationals within those flows. Secondly, more information on migration over the lifetime would be key to gain a comprehensive understanding of mobility patterns and tendencies within Europe. This connects to the third and final shortcoming of currently available statistics: the lack of data on circular and short-term migration, including cross-country commuting. Recommendations for bridging availability and quality gaps in the current state of European mobility data are offered throughout the paper.



Table of Contents

Introduction	3
I. Mapping strategy	4
II. The institutional framework and evolution of European migration data collection	7
III. Main sources of data on intra-European mobility	11
A. Types of data	11
<i>i. Types of administrative data</i>	<i>12</i>
<i>ii. Surveys</i>	<i>15</i>
B. Best datasets for researching intra-European mobility	16
<i>i. Databases based on administrative information</i>	<i>16</i>
<i>ii. Survey-based data</i>	<i>24</i>
IV. Available statistics and remaining challenges by thematic areas of EU mobility.....	31
A. Migration flows within the European Union	31
B. Stocks of intra-EU migrants	37
C. Statistics on selected aspects of intra-EU mobility.....	39
<i>i. Reasons driving migration within the EU.....</i>	<i>39</i>
<i>iii. Irregular migration.....</i>	<i>44</i>
<i>iv. Monitoring lifetime/multiple migrations</i>	<i>45</i>
<i>v. Short-term migration, circular migration, cross-border commuting</i>	<i>48</i>
Conclusion	51
References	53



Introduction

The first step in achieving an informed policy debate concerning the impacts and management of intra-European mobility is to develop a nuanced understanding of the scale and nature of EU mobility patterns. However, the complexities of migratory movements within the European Union can only be grasped with rich and accurate data tracking trajectories and the characteristics of movers. This aim of this paper is to provide a review of the existing data sources on European mobility; in doing so, we evaluate the degree to which the currently available body of data enables us to understand the phenomenon.

The importance of a regulated EU-level framework to produce reliable statistics on the continent's population – including its movements – was already recognized in the early 2000s, and led to a growing number of agreements aiming to establish a system of harmonized statistics at the European level, based on cooperation between all Member States. Indeed, the availability of basic indicators on stocks and flows of European migrants has improved vastly since the milestone 2007 Regulation. Nevertheless, considerable shortcomings remain concerning the depth of these figures: the aggregated nature of most available large-scale data (such as Eurostat statistics), for instance, allows for little nuance to be observed.

Therefore, a comprehensive review of existing data sources driven specifically by the objective of exploring the complexities of intra-EU migration may not only serve as a practical guide for EU researchers, but also as a call for action to policy-makers to improve the collection and publication of statistics related to European mobility.

Our desk review of databases has a deliberate focus on larger, cross-country data collections as the main sources of interest for their high representativeness and comparability at the regional level. Since the paper is thus not a classic inventory of data sources, in Section I the logic behind the strategy applied during our database-mapping exercise is presented. Next, Section II briefly outlines the evolution of the regulatory framework behind European statistics to provide a glimpse into the larger context



influencing the availability and quality of these data. The data themselves are discussed in a twofold approach: first, each main source is presented, including its strengths and weaknesses (Section III); second, the availability and quality of data is discussed by theme. Finally, the paper presents the authors' conclusions regarding the overall state of European mobility statistics, along with recommendations for improvement.

I. Mapping strategy

The goal of this mapping paper is to identify relevant databases and assess the degree to which they inform on intra-European mobility. Therefore, a natural first step during our mapping exercise is to define the information we are looking for in these data sources – as mentioned above, an in-depth understanding of mobility patterns demands more than just loosely defined basic migration figures. Indeed, to accurately map the scale and characteristics of intra-EU migration, the following information is needed:

- **Information on the individual mover's previous/next residence, citizenship and country of birth.** These three elements are crucial to define who we consider an *intra-EU* mover: someone who changed their residence from one EU country to another (regardless of their citizenship/country of birth)? Or only movements of the EU's own citizens (including naturalized residents)? Do we want to distinguish by origin as defined by country of birth? The size and characteristics of the measured population may vary substantially depending on what constitutes 'intra-EU'. Ideally, all three variables would be available for the same individual – or at least, the figure for each definition could be disaggregated by the other two. This would allow to identify relevant subgroups; for instance, the share of third-country nationals versus EU nationals within a specific country-to-country flow of migrants. While the differences in numbers between the three definitions may turn out to be relatively small (e.g., due to a low prevalence of secondary movements), the scale of the discrepancy cannot be known until these figures are available on a regional level.¹

¹ For a visual illustration of this issue (and further explanation), see Box 1 on page 35.



- **Aggregated values of flows for EU and non-EU groups** (following definitions above). Aside from simplicity in accessing EU-level mobility values at a glance, this is needed the availability because country-to-country level data is often imperfect, missing values for some or even most countries of origin or destination. As above, the simultaneous availability of figures as defined by all three terms (residence, citizenship, birth) would be ideal.
- **Migration history of the individual.** Information on multiple moves undertaken by individuals within the EU – or better yet, their migration trajectories over the entire lifetime – can shed light on a number of useful factors for better understanding EU mobility. Researchers could observe patterns of migration involving multiple movements, which could help assess whether and how certain first destinations are associated with specific second destinations, or predict migratory behaviour by assessing whether people who have once migrated are more likely to move again in their lifetime. Extending the period of observation can also help uncover long-term strategies of migration within the lifetime (for example, emigration during working-age years followed by return migration at retirement age), circular migration patterns, and information on length of residence(s).
- **Background information on movers.** This may include demographic (i.e., age, sex, marital status) and socioeconomic (employment, income, education/skills) information about migrant individuals, but also their motivation for migrating (work, education, family, lifestyle, etc.). Knowing these characteristics would help better understand the underlying factors influencing migration patterns – factors which risk being oversimplified when seen as merely a function of the general sending versus receiving environment. Understanding the composition of the migrant population in terms of these variables can help identify push and pull factors as well as preferred destinations of specific subgroups of migrants (defined, for example, by age and education level). In fact, it is possible that subgroups defined by such individual characteristics may show more coherent patterns of migration than groups defined by country of origin. Further, combining this information with data on sending/receiving countries can help identify the profiles of the typical movers for specific migration corridors.



- **Information on short-term, repeated movements.** Migration data typically focuses on movements that last at least one year and/or involve changes of residence, but this tends to exclude circular movements, including seasonal migration and cross-border commuting. Given the freedom of movement within the EU – together with other agreements and regulations facilitating the mobility of workers, students and others – these short-term, regular international movements have become a key aspect of everyday life in the EU, a phenomenon that would therefore be important to measure.

The above objectives set the basis for the review of migration databases; the paper examines the degree to which existing sources are able to meet these data requirements. In doing so, it identifies the most important sources and discusses their utility for research on intra-EU mobility, additionally pointing out remaining gaps in data. We would like to stress that providing an exhaustive inventory of all datasets containing some information on European migration (primary *or* secondary) is beyond the scope of this project.² Instead, the paper was shaped by the REMINDER project's aim to create a practical, goal-oriented “database of databases” for individuals interested in the research objectives above: we seek to identify the best available data for each objective, minimizing overlaps.

To achieve this, a macro-to-micro approach was taken. We performed a first, wide-ranging overview of existing sources of data based on well-known sources (e.g. Eurostat, OECD) and previous inventories of European migration data (Prominstat, THESIM); including a number of national level sources. Next, we identified the most efficient types of sources to include in our database and selected the main sources accordingly. After a detailed review of the availability and characteristics of the data available in our main – mostly macro – data sources, we identified remaining gaps and selected (or looked for) additional – typically more micro – data sources to fill these gaps.

² For these purposes, we recommend consulting completed migration data-gathering projects with an all-encompassing scope such as Prominstat A. Kraler and Reichel, *Statistics on Migration, Integration and Discrimination in Europe. PROMINSTAT Final Report*. (Brussels: European Commission, 2010). or THESIM Michel Poulain, Nicolas Perrin, and Ann Singleton, eds., *THESIM: Towards Harmonised European Statistics on International Migration* (Louvain-la-Neuve: Presses universitaires de Louvain, 2006).

Our criteria for choosing the “best” datasets were as follows: if multiple sources covering the same population, event or phenomenon were available, we included the most reliable and representative one; additional sources were only included if they contained complementary relevant information and therefore filled a gap in the coverage of time, geographic areas or variables of interest. Following this logic, databases sharing similar data that originates from the same sources – e.g. international databases publishing data from the same national administrative offices – were also considered overlapping and therefore only the most informative database was included (unless the databases somehow complemented each other). Furthermore, we aimed for up-to-date statistics, preferably no older than ten years but ideally available for the past decade to allow for trend observations. Comparability across countries was also a key asset, which put cross-country databases at an advantage. Given our focus, we targeted data on intra-EU movements in particular, but included more general EU migration data (including external movements) when data for the former was not available. Besides content, ease of use and accessibility were major considerations when including a dataset.

This strategy allowed us to create a comprehensive but efficient collection of the available data relevant to the research objectives listed at the start of this section.³ To further set up the basis for the paper’s mapping of datasets, the following section provides a short overview of the institutional context in which much of the discussed data is collected.

II. The institutional framework and evolution of European migration data collection

During the past decade, the EU has increasingly acknowledged the need for efficient collection of comprehensive migration-related information in order to develop effective migration policies. As a result, during the past few years, significant measures have been

³ The rationale for including certain datasets and excluding others from our inventory is further elaborated in Section III.



taken, in terms of both improving data collection practices and harmonising statistical data on the European level.⁴

One of the most important recent steps in this regard was the establishment of the European Migration Network (EMN) in 2002 with the aim of evaluating the quality of migration statistics across the EU. Similarly important was the Communication of the Commission regarding the 2003 Action Plan for the collection and analysis of migration statistics, which stressed data collection in areas such as: non-natural resident population, naturalisation, emigration and immigration, international protection, illegal entry and illegal stay, and residence permits of third country nationals. Until the late 2000s, however, much of the mentioned data was being collected on a voluntary basis. In addition, data collection was inconsistent and the level of harmonisation among different Member States was low.⁵

An increasing recognition for the need for more reliable EU-wide data led to the adoption of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection in 2007, a major milestone for the collection of migration-relevant statistics across the EU. According to the Regulation, Member States are required to provide Eurostat with statistics on the following:

- International migration flows, disaggregated by citizenship, country of birth, country of previous/next residence and by age and sex; population stocks disaggregated by citizenship and country of birth and by age and sex, acquisition of citizenship by country of previous citizenship;
- Asylum applications, decisions granting or withdrawing different forms of international protection status, asylum applications by unaccompanied minors, disaggregated by citizenship; and statistics on the operation of the Dublin III Regulation;
- Third country nationals refused entry to the Member State at the external border, third country nationals found to be illegally present under national immigration legislation, disaggregated by citizenship;

⁴ Kraler and Reichel, *Statistics on Migration, Integration and Discrimination in Europe. PROMINSTAT Final Report*.

⁵ Poulain, Perrin, and Singleton, *THESIM*.



- Residence permits issued to third country nationals, disaggregated by citizenship, length of permit validity and by the reason (immigration category) for the permit being issued;
- Third-country nationals subject to an order to leave the territory of the Member State under immigration legislation, and third-country nationals recorded as departing after the issue of such an order, disaggregated by citizenship.

According to the Regulation, the statistics should be based on sources such as records of administrative/judicial actions, registers of the population or relating to administrative actions, censuses, surveys, etc., according to their availability in respective Member States.⁶

Within the current institutional framework of migration data collection, the main providers of data are national statistical institutes (NSI) and relevant ministries of Member States. These organisations supply data to Eurostat, the statistical agency of the European Commission (EC). More specifically, statistics on migration flows, population stocks, and acquisition of citizenship are provided by NSIs. Statistics on asylum and on residence permits are provided by Ministries of Interior or related immigration agencies. Finally, statistics on the enforcement of immigration legislation are supplied by Ministries of Interior, immigration offices or Border Police. The Commission relies on these data during policy analysis and drafting of reports and proposals.⁷

Pursuant to the Regulation, reports on its implementation are published by the European Commission (EC) once in every three years, starting from the first report in 2012. Since the publication of the first report, several other legislative acts have been adopted to improve the consistency of statistics.^{8,9} The second report on the implementation of Regulation No

⁶ Council of the European Union and European Parliament, '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 (Text with EEA Relevance)', *Official Journal of the European Union*, no. L 199/23 (31 July 2007): 23–29.

⁷ European Commission, 'Report from the Commission to the European Parliament and the Council on the Implementation of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection' (Brussels: European Commission, 20 September 2012).

⁸ Regulation (EU) No 1260/2013 of the European Parliament aims at harmonisation on population-related data and sets out detailed characteristics of the required data, quality criteria, deadlines, etc., whereas Commission Implementing Regulation (EU) No 205/2014 of 4 March 2014 sets uniformed conditions for the



862/2007, published in 2015, notes that data collection in all areas has significantly improved in terms of accuracy, timeliness, and reliability.¹⁰ A significant improvement relates to consistency; prior to the adoption of the Regulation, inconsistency among statistical definitions used by different Member States was a major problem, leading to complications in terms of comparative analysis. Addressing this issue, the Regulation provides harmonised definitions, based on the statistical recommendations of the United Nations and relevant European legislation. According to the 2015 EC report, common definitions have contributed to increased comparability of data.

Overall, Regulation (EC) No 862/2007 has led to substantial developments in terms of migration data collection and analysis in the EU. Regular monitoring of the implementation of the Regulation can be expected to further refine the data collection system. In addition, the Commission regularly observes compliance with the Regulation and takes follow-up steps to address non-compliance by Member States.

implementation of Regulation (EU) No 1260/2013, specifically concerning international migration flows and population stocks disaggregated by citizenship and country of birth under Article 3 of the Regulation No 862/2007 [Council of the European Union and European Parliament, 'Regulation (EU) No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European Demographic Statistics Text with EEA Relevance', *Official Journal of the European Union*, no. L 330/39 (12 October 2013), <http://data.europa.eu/eli/reg/2013/1260/oj>; European Commission, 'Commission Implementing Regulation (EU) No 205/2014 of 4 March 2014 Laying down Uniformed Conditions for the Implementation of Regulation (EU) No 1260/2013 of the European Parliament and the Council on European Demographic Statistics, as Regards Breakdowns of Data, Deadlines and Data Revisions Text with EEA Relevance', *Official Journal of the European Union*, no. L 65/10 (3 May 2014): 10–26.].

⁹ In addition, four new legislative acts, relevant for compilation of statistics on residence permits under Article 6 of the Regulation No 862/2007 have been adopted [European Commission, 'Report from the Commission to the European Parliament and the Council on the Implementation of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection' (Brussels: European Commission, 30 July 2015).]

- Council Directive 2009/50/EC on the conditions of entry and residence of third-country nationals for the purposes of highly qualified employment;
- Directive 2011/98/EU of the European Parliament and of the Council on a single application procedure for a permit for third-country nationals to reside and work in a Member State and on a common set of rights for third-country workers legally residing in a member State;
- Directive 2014/36/EU of the European Parliament and of the Council on the conditions of entry and stay of third-country nationals for the purpose of employment as seasonal workers;
- Directive 2014/66/EU of the European Parliament and of the Council on the conditions of entry and residence of third-country nationals in the framework of an intra-corporate transfer.

¹⁰ European Commission.



Following this brief overview of the EU-level system of migration data collection, the next section discusses the sources of data we identified as most useful for researchers analysing mobility within the European Union.

III. Main sources of data on intra-European mobility

A. Types of data

A first distinction regarding the different types of migration data within the EU can be made between information originating from administrative sources versus sample surveys (compiled for statistical purposes). Given their differing strengths and limitations, the two types of sources may be best used to complement one another.

Administrative data is generally more representative, given that it aims to record the entire population or all events it targets (e.g. foreign-born nationals, or naturalisations), as opposed to sample-based efforts, which by nature work with estimations. Nevertheless, one should note that in particular cases a sophisticated sampling design may lead to a more exact estimation than a poorly implemented administrative recording system. Moreover, administrative records may exclude residents with an irregular status, while surveys that apply a sampling method not based on official records of inhabitants may be able to capture migrants who are ‘legally invisible’.

It is also important to recognise that although it is the primary source of many statistics, the system of administrative records is not typically designed to serve statistics.¹¹ For instance, the state might choose not to collect certain information about its inhabitants (or at least not record all personal information in one place), either because it does not see a need for it, or more deliberately, for privacy or political considerations. Therefore, only a limited range of variables may be available through individual data sources, which restricts not only the available information on mobility itself but also the options for disaggregating mobility data by multiple factors at once (which could allow an in-depth look at the different

¹¹ see also Kraler and Reichel, *Statistics on Migration, Integration and Discrimination in Europe. PROMINSTAT Final Report*.



demographic or socio-economic groups within a specific migration corridor, for example). Conversely, a key advantage of sample-based surveys is that they allow for the inclusion of a wider range of variables, tailored to help produce specific statistics.

An additional benefit of surveys is their relative flexibility compared to administrative data collection systems: changing the list of variables from one round of a survey to the next (e.g. to include questions on mobility) has a relatively low cost. Similarly, it is comparatively simpler to implement a cross-country survey using uniform methods than to convince countries to change their administrative systems in order to achieve harmonisation in regional statistics. Finally, panel surveys can provide a unique insight into trends and/or changes in migration-relevant variables over time in the individual's life, instead of a single snapshot.

In sum, administrative data provide a key basis of mobility figures in Europe, but this basis needs to be complemented with sample-based survey data for more specific and detailed information. Not all administrative data are created equal, however. Countries' registers, censuses, permits and enumerations of events (e.g. naturalisations) have different benefits and limitations in informing researchers about regional migration.

i. Types of administrative data

Population registers are typically considered the most reliable source of data on migrant stocks and flows, as – at least in theory – they register each incoming and outgoing settled inhabitant.¹² Their primary limitation is their imperfect availability, since not all European countries keep such registers; as of 2017, only 18 out of 28 Member States provided register-based data as a source for Eurostat's immigration statistics.¹³ Moreover, even when implemented, population registers do not always cover the entire target foreign population; for example, when foreigners are expected to present a residence and/or work permit valid

¹² Poulain, Perrin, and Singleton, *THESIM*.

¹³ Eurostat, 'International Migration Statistics Reference Metadata in Euro SDMX Metadata Structure (ESMS)', 2017, http://ec.europa.eu/eurostat/cache/metadata/en/migr_immi_esms.htm.



for at least as long as the minimum registration period to be recorded in the registry, non-complying (undocumented) immigrants are excluded.¹⁴

Deficiencies in registrations can pose a major challenge in tracking migration through population registers even among 'regular' migrants. Generally speaking, the willingness of inhabitants to report their movements depends on the advantages and disadvantages of being or not being registered, as well as the existence and the rigour of registration requirements. For these reasons, arrivals tend to be more thoroughly recorded than departures, for instance.¹⁵ Differences in registration criteria and practices across countries have impeded cross-country comparisons of migration statistics in Europe in the past. In recent years, however, the implementation of Regulation (EC) No 862/2007, together with a push for improvements articulated in the 2012 EC report and the following redevelopments of administrative systems have achieved greater homogeneity – or at least systematically documented deviations – in register-based migration data, as well as in migration statistics overall.¹⁶

Population censuses are another vastly comprehensive, and therefore highly representative source of information on immigration across Europe. Besides not recording emigration, their main limitation as a source of mobility data is that they are generally conducted only every five to ten years, leaving a gap in data for the time periods between censuses (also called *intracensal* years). A number of countries therefore use census data in combination with other sources (e.g. registers or surveys), for example complementing census data with other data in intracensal years, or using census data to periodically revise data collected in other ways.¹⁷ Although they do not cover people not living in private households (e.g. reception centres, hostels for immigrants), an advantage of censuses is that they may be able to capture a portion of the irregularly residing population excluded from population

¹⁴ OECD, 'Statistical Annex', in *International Migration Outlook 2016* (OECD Publishing, 2016), http://dx.doi.org/10.1787/migr_outlook-2016-46-en.

¹⁵ Kraler and Reichel, *Statistics on Migration, Integration and Discrimination in Europe. PROMINSTAT Final Report*; OECD, 'Statistical Annex'.

¹⁶ European Commission, 'Report from the Commission to the European Parliament and the Council on the Implementation of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection', 30 July 2015.

¹⁷ Poulain, Perrin, and Singleton, *THESIM*.



registers and residence permit records (ibid.). Still – similarly to registers – censuses are most useful for basic immigration statistics since they only accommodate few questions.

In terms of coherence and accessibility, the EU-wide 2011 Population and Housing Census marked a remarkable step forward: for the first time, data from censuses conducted in the Member States was produced following European legislation ¹⁸ to ensure quality and uniform, comparable outputs. EU-wide census rounds were also conducted in 2001 and 1991, but these were merely based on ‘gentlemen’s agreements’ instead of EU legislation.¹⁹ The harmonised Census 2011 data is available on Eurostat’s general database as well as through the Census Hub, a new transmission system developed specifically to disseminate the 2011 Census.

In addition to population registers and censuses, administrative records targeting specific subsets of the population comprise an important source of data on mobility within Europe. Statistics related to naturalisations, asylum, residence permits and the enforcement of immigrant legislation, for instance, are often produced through administrative records collected by specialised agencies within the ministries of interior, related immigration agencies, or border police. Moreover, the use of data from specific registers such as health insurance registers or tax registers has been noted as a key step in improving European migration statistics in the Commission’s 2015 Report.²⁰

Some typically rich sources of information for migration data are, however, less informative when solely focusing on intra-European movements. Residence permit records, for instance, are usually a useful source because they contain information on migrants’ (declared, planned) length of stay and reason for migrating. However, regarding the subject of intra-European movements this data is of limited use: it excludes EU nationals (since they do not require a permit to stay), and does not track the previous country of residence, so the only

¹⁸ Regulation (EC) No 763/2008, Regulation (EC) No 1201/2009, Regulation (EU) No 519/2010, Regulation (EU) No 1151/2010

¹⁹ Eurostat, ‘Census 2011 Round (Cens_11r) Reference Metadata in Euro SDMX Metadata Structure (ESMS)’, 2014, http://ec.europa.eu/eurostat/cache/metadata/en/cens_11r_esmscs.htm.

²⁰ European Commission, ‘Report from the Commission to the European Parliament and the Council on the Implementation of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection’, 30 July 2015.



“origin” country recorded for covered third-country nationals is their country of citizenship – even if they are in fact migrating from another EU country.

A focus on migration from external borders is also evident in asylum-related data. Since the country considered responsible for the asylum procedure under the Dublin regulation is usually the country through which the asylum-seeker first entered the EU, most of the asylum-related data tracks arrivals from outside the borders of Europe. The exception that is relevant intra-European movements is data on ‘Dublin transfers’, which refers to asylum-seekers whose asylum procedure is being transferred from one Member State to another following a request for either one country to ‘take charge’ of the application instead of the original country assigned, or to ‘take back’ the asylum-seeker from another country to the original country assigned.²¹

ii. Surveys

The role of sample-based surveys in migration data is twofold: firstly, they are uniquely able to add nuance to the basic migration statistics provided by administrative records; secondly, they often serve as complements or alternatives to administrative data when it is incomplete or missing altogether. Focusing on intra-EU mobility data, we distinguish between cross-national and national surveys. National gaps in administrative data collection are typically filled with the use of national surveys (e.g. UK International Passenger Survey for the UK’s migrant flows statistics). Concerning surveys as sources for more nuanced data, however, cross-national surveys are more practical tools than national surveys for research spanning multiple European countries: similar sampling procedures, definitions and overall homogenous surveying methods allow for the data from different countries to be used as one coherent dataset. The benefit of national surveys, besides being the only available option in some cases, is that they are more likely to offer panel data than cross-national surveys and can target questions relevant to the specific country context; this makes them a useful (additional) source even when cross-country surveys are available.

²¹ Eurostat, “‘Dublin’ Statistics (Migr_dub) Reference Metadata in Euro SDMX Metadata Structure (ESMS)”, 2010, http://ec.europa.eu/eurostat/cache/metadata/en/migr_dub_esms.htm.



For both types of surveys, the most critical limitation of existing data is the lack of migration-specific surveys. Even though questions enabling the identification of migrant individuals are often included (e.g. country of birth), the lack of an oversample of migrant individuals makes their subsample too small for analysis. Furthermore, ‘migration-specific’ surveys usually target the topic of integration, focusing on current socio-economic indicators of the migrant and overlooking questions related to the act of migration itself (i.e. previous movements, motivations for doing so, plans to return etc.).

B. Best datasets for researching intra-European mobility

i. Databases based on administrative information

a. International online (administrative) databases

We find that the most practical way to access large-scale administrative data on intra-EU mobility today is through the online international databases which collect data directly from Member States’ national statistical institutes and share that data on an online platform in a uniform system. International databases have multiple advantages for researchers, particularly for those with an international focus: these platforms provide open access to a comprehensive collection of European migration-relevant data in English, organised in a manner that allows for cross-country comparisons; they conveniently present multiple Member States’ data all in one place in the same format, broken down by topics, years, and other available disaggregations. The metadata included with the datasets typically contains information on the primary sources of the data and notes differences in methodology which further supports comparability.

Such databases are thus generally the most user-friendly and efficient source for comparative European migration statistics, especially when trying to gather cross-country administrative data. It is important to note, however, that the presented data is typically in aggregated form and has a limited variable list. In case of in-depth country studies, it is therefore advisable to also consult the NSIs directly, as they might be able to provide further datasets or more detailed versions of the data.



Eurostat database

Eurostat offers the most comprehensive database of European statistics overall, including data on migration within the region. As explained in detail in Section I, Eurostat regularly collects a variety of mobility-related statistics directly from the NSIs within a harmonised framework.²² This data is then shared on a freely accessible online database on Eurostat's website, which presents the information categorised by themes. Under each theme, a number of interactive tables present subsets of the data including all Member States plus EFTA countries. The thematic data is often split into sub-themes and different disaggregations (e.g. citizenship, age, sex) across a list of tables.²³ This system allows for a practical overview of figures for all countries, with a manageable amount of data per table.

The drawback of splitting up the available disaggregations of the data into separate tables is that it limits opportunities for the cross-tabulation of information. For example, immigrant flows may be disaggregated by citizenship, country of birth or previous country of residence, but not any two at the same time;²⁴ this makes it impossible to assess the volume of specific subgroups (see Section I). Furthermore, given that the data originates primarily from administrative records, the overall list of available background variables of individuals is limited (typically to age/age group, sex, country of birth, and citizenship).

The quality of produced statistics is ensured through technical guidelines (e.g. harmonised of definitions and preferred data sources) and validation checks. The benefits of this regulated data gathering process are reflected in the relatively high availability and quality of Eurostat migration statistics, which cover most European countries for the past decade²⁵

²² Besides the mobility-related datasets mentioned in Section I (categorized under the *Population* and the *Asylum and Managed migration* themes in the Eurostat database), under the *Education and training* theme Eurostat's online database also publishes the joint UNESCO-OECD-Eurostat (UOE) database on education which includes key data on student mobility within Europe (for more details, see Section III).

²³ Also worth noting is the previously mentioned Census Hub within Eurostat, an online platform specifically designed to allow national statistical institutes to share census data directly with users.

²⁴ The exception is available the cross-tabulation by citizenship and broad group of country of birth (and vice versa), where broad group indicates EU vs non-EU etc. While this is relevant, a much more useful cross-tabulation would include the country of previous residence to capture the subgroup of third-country nationals moving within the EU and exclude EU citizens arriving from third countries.

²⁵ Some data is available even further back: total migrant flow and stock data, for example, are available starting 1998 and 1999, respectively.



– with some exceptions. Gaps are notable especially regarding the more detailed data for specific countries. For instance, total immigration flow figures are available for almost all countries for the past ten years; however, a sub-selection of immigrants with a previous residence in an EU-28 country is only available starting 2013, while data specific to individual countries of previous residence continues to be missing for nearly half of the Member States.

Another important advantage is the availability of metadata for each dataset, even if the quality of these documents varies. A fairly common issue is the lack of indication of exact data sources for individual countries; sometimes completely sources are missing altogether, in other cases blanket terms such as “administrative sources” are used. It would also be helpful to know why data are missing for specific countries and/or years within certain themes, i.e. whether the country systemically did/does not collect the data, did not collect it for a specific period, chose not to share the data, or there was a problem with the quality of the data. In general, knowing in which cases the data is collected but not shared versus it not being collected altogether could help estimate chances for future (or conditional) availability of data. On a positive note, precise definitions (including occasional differences in definitions) are always specified in the metadata, which is a great aid for cross-country research.

An additional tool that makes Eurostat databases fit for cross-country research is the availability of variables coded following international classifications (e.g. broad citizenship groups, occupation, education). Finally, a key benefit of the Eurostat database is its publicly accessible online *bulk download facility*, which enables users to easily download multiple datasets at once; multiple formats are available, allowing to work with the data in different types of data analysis software.²⁶

All in all, while Eurostat is the most practical first source for a comprehensive overview of EU-focused mobility statistics, due to remaining gaps in the availability of data for particular

²⁶ The bulk download facility contains files in XML (Extensible Markup Language) or TSV. Further formats including XLS, CSV, SPSS, or PDF are available through download in the data explorer tool.

countries – further explored in Section IV – the database has not yet reached its full potential. For the full picture, researchers may need to resort to other sources, including UNDESA and OECD or the relevant NSIs, or, depending on the topic, surveys such as LFS or the Eurobarometer.

United Nations Global Migration Database (UNDESA)

The migration figures produced by the United Nations Department of Economic and Social Affairs (UNDESA) are perhaps the most widely used, on a global level. Although the main asset of UNDESA data are their global (or at least multi-continental) scope rather than their depth, they comprise a valuable back-up to Eurostat data. The latest – 2015 – revisions for both international migrant flows and stocks are directly available online. Generally speaking, UNDESA’s database on migrant flows contains more useful for our purposes than the migrant stock data collection. A key advantage of both is that they cover all 28 Member States (providing comparable data in a comprehensive collection), aim to capture all bi-lateral (country-to-country) migration, and are accompanied by good quality metadata. Drawbacks include a lack of aggregated EU-level figures (although ‘Europe’ as a region may be present), the format of the data, and missing data. Despite these shortcomings, UNDESA statistics may prove useful to fill in country/year specific gaps in Eurostat data. Advantages and disadvantages of the two UNDESA databases are discussed in more detail in Section IV.

OECD Migration Data

The Organisation for Economic Co-operation and Development (OECD)’s International Migration Database is another major source of migration-relevant information obtained directly from national correspondents. Covered topics include inflows and outflows of foreign population, inflows of asylum seekers, stocks of foreign and foreign-born populations and labour, and acquisitions of citizenship. From an intra-EU mobility perspective, however, this platform is of limited use. Firstly, OECD migration data in general excludes the six non-OECD member European Union states (Bulgaria, Croatia, Cyprus, Lithuania, Malta, and Romania). Secondly, disaggregations are only available by country of nationality or country of birth and sex.



The coverage is therefore limited compared to Eurostat (and UNDESA) in terms of geography, themes and background variables. Nevertheless, within the covered areas, the OECD's database occasionally contains data that it is noted as missing in Eurostat and/or UNDESA databases, which makes it a potentially useful complementary source. Like in the case of UNDESA, an additional benefit of the OECD database is the detailed metadata available for type of migration statistics, which is often more thorough in reporting the respective sources of data for each country than Eurostat. Also worth noting is the 2016 International Migration Outlook's Statistical Annex,²⁷ which provides an overview of recent OECD migration statistics by theme and country (including rich metadata).

Other international databases

Multiple other popular international databases collect large-scale (mostly) administrative data relevant to migration. When it comes to intra-EU migration in recent years, however, these datasets hold little to no additional value compared to Eurostat (especially once complemented with UNDESA and/or OECD data).

The Global Migrant Origin Database (DRCM) and IPUMS, for instance, are both impressive global census data collection efforts, but face limitations: data from the former is from around the year 2000, while the latter only contains half of EU countries. Overall, Eurostat provides better access to EU countries' census data (up-to-date and for all Member States), either through its general database or the Census Hub.

A number of other popular migration databases, such as UNICEF Data or MPI's data hub, essentially just reproduce data from one or both of two sources: UNDESA and UNHCR.

UNDESA data is best accessed directly; UNHCR provides wide-ranging data on asylum-related migration and could in fact be useful for complementing migration data for countries where asylum-seekers are not included in immigrant stocks or flows. When focusing on intra-EU movements, however, Dublin transfers and returns are the only

²⁷ OECD, 'Statistical Annex'.



relevant portion of official asylum-seeker data, which are already made available by Eurostat.

Further popular sources for migration statistics in general include the International Labour Organisation (ILO) and the World Bank, but again, these are not informative sources for recent data on intra-EU movements. ILO statistics do not include EU countries. The World Bank migration data includes net migration indicators (already available through Eurostat) and the Databank on Global Bilateral Migration – which could be a very useful tool if continued, but currently stops at the year 2000.²⁸

b. National statistical institutes as complementary data sources

Moving on to national statistical institutes (NSIs) as complementary sources of statistics, we find that while the data provided by individual NSIs might be richer than that found in international compilations, the NSI approach is often left efficient – especially when trying to collect comparable data on multiple countries. One of the main drawbacks of obtaining data from the NSIs is that it often involves a slower, multiple-step process. Access might be limited to nationals or to on-site use, and datasets may only be available in the local language. Furthermore, the researcher may not benefit from the homogeneity (in methodology, definitions, variables) of the data that countries are required to provide when the information is collected within a single overarching framework.

This sub-section explores how the data disseminated by the individual national statistical institutes compare to those available via the Eurostat database, noting some regional trends in the quality and accessibility of data. Information is based upon a content analysis of each website of the various statistical institutes of the EU Member States.

With a few exceptions, migration data at the national level are generally less accessible and more scattered as compared to Eurostat. While most states have information on the country of origin and citizenship of migrants, data on country of previous residence are publicly available in only 13 member states (note: this about the same as Eurostat). At the

²⁸ The World Bank also publishes information on remittance flows (and how they compare to migrant flows), which falls outside the focus of this project but is noteworthy for migration research in general.



national level, there is a tendency to aggregate immigration data by macro-area of origin (EU, non-EU), which constrains the full potential of those statistics. This trend is particularly visible in Eastern European countries such as Hungary, Poland, Estonia, Croatia, and Lithuania, but also in smaller states such as Malta and Portugal. In addition, these countries, as well as Southern European countries, France, Belgium, and Luxembourg only translate partial information to English, limiting the international accessibility of their data.

As mentioned above, we note a tendency among most Eastern European countries, small countries and countries characterized by lower economic performance to provide less exhaustive migration data compared to their larger and/or wealthier peers. This group includes **Croatia, Slovenia, Hungary, the Czech Republic, Slovakia, Poland, Lithuania, Estonia, Luxembourg, Portugal, Cyprus, Malta, Romania and Bulgaria**. Migration statistics of these countries are generally speaking less available at the national level than at the supranational level (Eurostat, OECD, and UNDESA have more information), as data are not always reported and/or open-access.

Despite this shared trait, this remains a highly heterogeneous group. Some countries,²⁹ such as Slovakia, have no migration section on their website, nor migration data. Croatia provides information on country of birth and place of residence in Croatia, but no information on age or gender. Similarly, Luxembourg only shares the aggregated the number of immigrants. Other countries, such as Romania, Belgium, and Poland, also have an aggregated figure for immigrants (or ‘foreigners’), but this can be disaggregated by age and gender. Finally, some countries indicate countries of origin as well as age and sex of migrants (Estonia). Other provide information on migration, but it is partial: Luxembourg, for instance, only shares aggregated figure of immigrants; Croatia provides information on country of birth, but not age or sex; Romania, Belgium, and Poland also aggregate immigrants by a single definition without differentiating by origin countries, but include and age and sex disaggregation;

²⁹ Note: by *countries* here we refer to the respective countries’ national statistical institutes (more specifically, the information available on their websites). NSIs were also contacted individually via email for more accurate information, but not all contacted institutes have responded; the information learnt as a result of this effort has helped complement the findings presented in this section.



finally, some countries indicate all three – country of origin, sex, and age – but this information is still poorer than that found in the Eurostat database.

A suggestion for these countries would be to follow the example of the Maltese National Statistical Office. On their website, they have a section dedicated to EU statistics, in which they link the user to international databases in which he can find more information. This kind of outsourcing mechanism could be easily implemented by other offices, and it would result in a complete picture of the existing statistics of a given country.

Austria, Finland, France, Greece, Italy, Latvia, Netherlands, and Spain represent a second group of countries. In this group, migration statistics at the national level are very similar, if not equal in terms of data available to the statistics that can be retrieved from supranational databases. Almost all of these countries' NSIs provide statistics on the country of birth, country of citizenship, country of previous residence, age, sex, education, occupation, and marital status of the immigrant population. These statistics are generally available from early 2000 until 2016. The main limitation of the statistics provided and disseminated by those countries – except for the Netherlands and Finland – is that yearbooks, reports, and articles are only partially retrievable in English. Statistics in Latvia and Greece are largely based on the 2011 Population and Housing Census so most of them are only available since 2011. In Spain and Italy, microdata from administrative sources integrate information on migration and sometimes represent unique information that is not reported on Eurostat. In Greece, by looking at output tables and reports of the Hellenic Statistical Authority, it stands out that with the full census data, it will be possible to cross-tabulate between country of origin, citizenship, and country of previous residence. This is also possible with Austrian statistics. Cross-tabulation represents a unique feature, which is not captured by Eurostat statistics, yet it is crucial to understand intra-EU mobility.

The last group of countries is composed of **Denmark, Germany, Ireland, Sweden, and the United Kingdom**. These countries have more exhaustive migration statistics at the national level than at the international one. Statistics are by large open-access and translated into English. Statistics of the United Kingdom that might be considered a 'good practice' include



information on occupation prior to migration, reasons for migration, and previous reasons for migration – it should be noted, however, that these statistics are largely survey-based (UK International Passenger Survey), and therefore limited in accuracy. In Germany, the Research Data Centre (FDZ) of the German Federal Employment Agency (BA), has a variety of microdata originating from administrative registers and surveys with very accurate and vast information on migration and labour market history variables. This data is highly valuable since in-depth German migration data is at times missing from Eurostat database. Ireland has information on the level of English, religion, level of education, ethnicity, and field of study of migrants. Swedish data can date back to 1968 and have information on seasonal and circular migration. In the United Kingdom, Ireland, and Denmark it is also possible to cross-tabulate between the country of origin, citizenship, and country of previous residence.

All in all, the degree to which the administrative data available through NSIs can be used to complement Eurostat data is highly dependent on the individual country: as presented above, cases vary from entirely missing to very rich data.

ii. Survey-based data

a. Cross-national surveys

As previously mentioned, for the purpose of studying intra-EU mobility we find larger, international surveys to be preferable to smaller-scale, national surveys, given their advantage in representativeness and comparability of cross-country results. An overall larger sample size is beneficial because surveys that would otherwise address relevant topics often lack a significant oversample of migrants. The ‘ideal’ intra-EU mobility survey would in fact be have an EU-wide coverage with a migrant oversample – or exclusively migrant sample – in each country, with questions addressed specifically at migratory behaviour (including migration history) and relevant background variables capturing (also) individual characteristics the time of migration. This would help see EU mobility decisions in the personal context of individuals, enabling a better understanding of the drivers behind observed patterns.



At the time of this report, such a survey had not been created. We therefore discuss the existing surveys that best approximate the set of features outlined above: most notably, the European Internal Movers Social Survey (a one-time survey from 2004), the Labour Force Survey, and selected waves of the Eurobarometer.

Labour Force Survey

The Labour Force Survey (LFS) may be considered the most useful on-going survey for statistics on intra-European mobility. The LFS is the largest European household sample survey, producing comparable data across all 28 Member States (plus two candidate countries and three countries of the European Free Trade Association).³⁰

The core questionnaire of the LFS collects quarterly data on labour participation of Europeans aged 15 and over. It allows for the identification of migrants through their country of birth and nationality, but with limitations: countries of nationality and birth are aggregated into the following groups: national/native; EU15; 2004-accession countries (10 in total); 2007 and 2013 accession countries (3 in total); EFTA; other Europe; and groups for other main regions of the world outside Europe. This limitation prevents researchers from using LFS data to identify bilateral country corridors and the obviously useful analysis of migrants with specific countries of origin. Nevertheless, given that the EU accession timeline followed some regional patterns, the grouping does at least allow to distinguish between some main regions within Europe. Years of residence in the surveyed country are also included, which is fairly rare information for migrants who are EU nationals. Besides a number of labour and socio-economic characteristics, migration-relevant variables include the country of place of work (which enables to identify cross-country commuters), the year when the highest education was obtained (comparable to years of residence to see if it was obtained in the host country), and the country of residence one year before the survey, which could shed light on the trajectory of (likely) the most recent migration. However,

³⁰ The quarterly LFS sample size across the EU was about 1.6 millions of individuals in 2015 [Eurostat, 'European Union Labour Force Survey: Description of the Dataset', 2017, <http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey>.]



despite the large size of the overall sample, the relatively small size of the migrant subsample limits the benefits of the LFS.³¹

In 2008, the annual ad-hoc module focused on the labour situation of migrants and their descendants, oversampling migrants, for a total of ca. 1.44 million observations³². A similar module was carried out in 2014 and is planned again for 2021. The ad-hoc modules cover most EU-28 countries, with the exception of Croatia and Finland in the 2008 round as well as Denmark, Ireland and the Netherlands in the 2014 round; furthermore, access to microdata is not available for Germany.

Regarding information on intra-European mobility, two variables stand out from the 2014 ad-hoc questionnaire: *Reason for migration* and *Last country of work abroad*. The former targets a key gap in knowledge regarding the drivers of intra-EU migration, which is especially difficult to track for those who can move freely within Europe (residence permit data typically includes the nature of migration). Information on the last country of work abroad can be used to track potential multiple movements across Europe.

Both the core of the thematic ad-hoc modules of the LFS are therefore useful additional sources for information on European mobility, but with room for improvement. Assuming that an oversample of migrant respondents is unlikely to be implemented in the core survey – although this would be most helpful – improvements in mobility measurement could be achieved with the addition of the following variables to the upcoming ad-hoc modules: planned length of stay, future migration/return plans, previous countries of residence in Europe (incl. years), and labour and education characteristics at time of migration.

From a practical point of view, difficulty of access is an issue when it comes to mobility-relevant data within LFS. The aggregated data available through the Eurostat online

³¹ Eurofound, 'Analysis of the Socioeconomic Situation of Migrants – Gathering Comparable Data on Their Living Conditions', Workshop report (Brussels: Eurofound, 17 November 2010), http://csdle.lex.unict.it/Archive/LW/Data%20reports%20and%20studies/Reports%20and%20studies%20from%20EUROFOUND/20110325-115319_Eurofound_Migrants_March11pdf.pdf.

³² H. Lien and I. Toleikyte, 'The Labour Market Situation of Migrants and Their Immediate Descendants Evaluation of the 2014 Labour Force Survey Ad Hoc Module' (Luxembourg: Eurostat, December 2015), http://ec.europa.eu/eurostat/documents/1978984/6037334/Evaluation_report_AHM_2014.pdf.



database do not include those variables that are the most useful sources of additional information compared to administrative data (in terms of years of residence, prior migrations, reasons for migration and cross-country commuting). Microdata including all variables is accessible for researchers, but subject to a fairly complex and lengthy (8-10 weeks) application process. In addition, Germany does not provide access to its LFS microdata – a considerable impediment to our type of research, given that the country is both a major receiver and sender of EU movers.

European Internal Movers Social Survey (EIMSS)

The European Internal Movers Social Survey, carried out in 2004, was the first – and, to our knowledge, only – large-scale systematic survey-based study of intra-EU migrants. EIMSS was carried out as part of the PIONEUR project, which aimed to fill the gap in knowledge about the socio-demographic profile and, in particular, the motivation, life patterns and personal consequences of migration for European citizens who have migrated from one Member State to another. The sample contained 5,000 European citizens residing as foreigners in France, Germany, the United Kingdom, Italy, and Spain.³³ The dataset is highly informative since it is one of the very few sources containing detailed information on lifetime migration within Europe (asking every other country the respondent has lived in, prior migration to the current destination country, reasons for settlement, and future moving aspirations including retirement), among other relevant variables. The main drawback of the survey is that it was carried out over a decade ago; a follow-up round (ideally extended to include more European countries) could provide invaluable information on the patterns and drivers of intra-EU mobility. Access is extremely easy, immediate and at no cost: the primary data is available to download online for researchers via the GESIS database.³⁴

³³ Ettore Recchi et al., 'Geographical and Job Mobility in the EU (Final Report)', 'Empirical Evidence on Job and Geographical Mobility in the European Union' Tender No. VT/2005/0107 DG Employment, Social Affairs and Equal Opportunities (CIUSPO, University of Florence, June 2006).

³⁴ <https://dbk.gesis.org/dbksearch/sdesc2.asp?no=4512>



Eurobarometer

The Eurobarometer is primarily known as the EU's main public opinion survey, but some of its waves have actually targeted key questions related to understanding European mobility. Specifically, the 2005 Eurobarometer survey on *Geographical and labour market mobility* (wave 64.1), the 2007 Eurobarometer on the *Geographical mobility of citizens*, and the 2009 follow-up to the former (*Geographical and labour market mobility*, wave 72.5) addressed topics including Europeans' mobility experiences and intentions, as well as reasons encouraging or discouraging people from moving. For instance, topics covered in the 2009 Eurobarometer include:

- Respondent's opinions about the impact of people moving across regions or countries within the EU on individuals, families, the economy, the labour market and European integration.
- Respondents' experiences of living, working and/or studying abroad
- Respondents' plans to work abroad in the future
- Positive and negative experiences of those who have already worked abroad
- Motivations and disincentives for working abroad
- Perceived issues to be faced when working abroad
- Ways respondents think they would find work abroad
- Respondents' knowledge of EURES (European Employment Services), and the services they would look for in an employment service.³⁵

Given the sample size of 1,000 interviews, it should be noted this data serve not so much as a source for exact figures on migration flows within the EU, but rather as an insight into mobility tendencies and intentions, as well as the influencing factors that shape these attitudes and decisions. Additionally, the 2009 data is becoming relatively old: a new – post-financial crisis – wave of the special survey on mobility would be highly useful. Like EIMSS,

³⁵ European Commission, 'Geographical and Labour Market Mobility', Special Eurobarometer 337 (European Commission, June 2010), 5–6, <http://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/803>.



primary data from the Eurobarometer surveys is very easy to access via the GESIS online database, where users can freely download the data in formats compatible with multiple types of data analysis software.³⁶

Other international surveys

Besides LFS, Eurostat's other major survey gathering information on the socio-economic situation of Europeans is the European Union Statistics on Income and Living Conditions (EU-SILC). It is similar to the LFS design and implementation, but focuses more on income and issues related to social inclusion and poverty. It also allows to identify migrants through country of birth and citizenship and years of residence (latter two not included for all countries), and provides information on the socio-economic background of individuals. However, it works with a significantly smaller sample than LFS³⁷ – making the absence of a migrant oversample a more severe problem – and lacks some of the useful additional mobility-related variables included in LFS (e.g. information on commuters, country of residence one year prior to the survey). In summary, given EU-SILC's similarities but lack of additional value, we consider LFS a preferable source of survey data to EU-SILC for researching intra-EU migration.

Similarly to the Eurobarometer, the European Social Survey is a major cross-country survey focusing on attitudes and opinions of Europeans. However, contrary to the Eurobarometer, for the ESS no waves targeting EU mobility have been implemented; the latest wave focusing on attitudes towards immigrants was carried out in 2014, but only one question concerned EU migrants, asking respondents to what degree they would allow immigrants from poorer European immigrants into the country.³⁸ Thus, we do not consider the ESS a very useful source for understanding European migration.

³⁶ <https://www.gesis.org/eurobarometer-data-service/search-data-access/data-access/>

³⁷ EU-SILC's minimum effective sample size is below 300,000 individuals in total [Eurostat, 'EU Statistics on Income and Living Conditions (EU-SILC) Methodology – Sampling', Eurostat Statistics Explained, 2016, [http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_\(EU-SILC\)_methodology_-_sampling#Sample_size.](http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_sampling#Sample_size.)]

³⁸ European Social Survey, 'Attitudes towards Immigration and Their Antecedents - Question Design Final Module in Template' (Centre for Comparative Social Surveys, City University London, 2015),



International student or graduate surveys can be relevant even if they do not specifically focus on migration because they provide additional insight into a specific (and young) cluster of the population. The OECD's Programme for International Student Assessment (PISA) is a triennial survey testing the skills and knowledge of 15-year-old students worldwide, including all EU countries in its latest (2015) round. The list of migration-related variables is limited (country of birth, years of residence), but PISA nonetheless can be a helpful additional source of data on the adolescent (15-year-old) cohort within migrants across the EU. PISA data is freely available through its online database.

REFLEX and its extension for Eastern Europe, HEGESCO, are two large-scale European surveys for higher education graduates. Altogether the two projects cover 18 European (16 EU-28) countries.³⁹ The surveys did not oversample migrants and the data is somewhat old (from 2005 and 2007, respectively, targeting graduates from 5 years earlier); it is, nevertheless, highly valuable since it can provide rare insights on extended migration trajectories of graduates, by specifying their country of birth, country of residence at age 16, during higher education studies, when first starting employment, and at the time of the surveys. The datasets are freely available for research purposes via request by email.

b. National surveys

We find few national surveys that provide useful additional nuances on intra-EU migration considering the information available through Eurostat and LFS. Three national surveys that do stand out are the German Socioeconomic Panel (SOEP) and the United Kingdom's (UK) Understanding Society Survey and International Passenger Survey.⁴⁰ SOEP is a remarkably rich source of data as it looks into lifetime migration with detailed questions regarding past and future migrations; the panel nature of the data further helps to track migration trajectories. The data available through SOEP is particularly important given the lack of

http://www.europeansocialsurvey.org/docs/round7/questionnaire/ESS7_immigration_final_module_template.pdf.

³⁹ Austria, Finland, France, Germany, Italy, the Netherlands, Norway, Spain, UK, Belgium-Flanders, Czech Republic, Portugal, Switzerland, Estonia, Lithuania, Poland, Hungary, Slovenia (plus some non-European countries).

⁴⁰ Ireland's Quarterly National Household Survey is also a rich source of national data, but for the purpose of this paper we regard it as a part of the Labour Force Survey and therefore do not discuss it separately



access to microdata for Germany in the LFS and gaps in the data shared by the country via Eurostat. SOEP microdata are freely available to researchers (upon request via email).

Selected waves of the UK's Understanding Society survey comprise a similarly valuable – and rare – source of data on migration history: wave 1 (2008) asks both natives and immigrants detailed information on the countries they have lived in prior to the time of the survey. These include the number of countries and the location (up to 5), questions on moves before and after having first moved to the UK, and age when moved to the UK. Additionally, it includes questions on internal migration. Wave 3 in 2010 only had two questions on future plans to migrate (similar to those in Eurostat). Wave 7 (2014), on the other hand, repeated not only all variables from wave 1, but also added reason of migration and current migration intentions (yes or no) to the variable list.

Lastly, the UK's International Passenger Survey (IPS) includes information on reason for migration, country of previous residence, usual occupation prior to migration, and immigrants' previous stay in the UK, in addition to tracking short-term and tourism-related mobility. Datasets derived from both the Understanding Society survey and IPS are available through the website of the UK Data Service portal (upon registration).

Following this review of the main sources on data relevant to intra-EU mobility, the next section discusses the quality and availability of statistics by thematic areas within the study of intra-EU mobility and outlines remaining gaps in the currently available body of data.

IV. Available statistics and remaining challenges by thematic areas of EU mobility

A. Migration flows within the European Union

Data on intra-European migration flows is, in theory, easily available through administrative records of migrants' (de-)registrations when they change their residency. Eurostat publishes this data (as forwarded by NSIs) in a comprehensive database with fairly well comparable



numbers.⁴¹ However, the accuracy of these figures is challenged for both EU citizens and third-country nationals.

Concerning EU nationals, the problem is that a number of intra-EU movements of EU citizens are undetected by administrative records: registrations and de-registrations are often voluntary in nature and EU citizens do not require permits to reside in another Member State. As mentioned in Section III, emigrations in particular tend to be underreported because there are neither strict regulations nor benefits encouraging emigrants to de-register (as opposed to the case of registering in a new country).⁴² This weakness of emigration statistics can be somewhat be solved using mirror statistics, the quality of which depends then on the receiving countries.

In the case of third-country nationals (TCNs), entry into the EU – if legal – is well documented through residence permits (in addition to registers). The problem is differentiating between these “external” movements into the EU and subsequent “internal” movements that happen across Member States. The most efficient way to capture intra-EU migrations of TCNs would be through their country of previous residence, which is in fact mostly recorded in Eurostat flow data (discussed further below). When this information is lacking, however, there is a risk of incorrectly assuming the country of birth or citizenship is the starting point of the studied mobility flow for TCNs, leading to an underestimation of TNC mobility within the borders of the EU.⁴³ The issue of imperfectly overlapping definitions and multiple disaggregation is further explained in Box 1 on page 35.

⁴¹In Eurostat data, the definition of immigration is consistently based on change of usual residence for at least 12 months (although minor differences continue to exist in whether this refers the actual or the intended stay). A persisting source of heterogeneity regards the inclusion of asylum-seekers in migrant flow statistics (12 of the EU-28 countries include them, 16 do not) [Eurostat, ‘International Migration Statistics Reference Metadata in Euro SDMX Metadata Structure (ESMS)’]. Most countries base their data on registers and/or other administrative sources (e.g. censuses, residence permits), while a handful of countries rely on survey-based methods (often combined with census data) [ibid.]. Overall, the comparability of European data is still imperfect; however, the fact that these remaining differences and gaps are now fairly well documented can enable researchers to deal with them. For more details on Eurostat’s database, see the previous section (III).

⁴² OECD, ‘Statistical Annex’.

⁴³ For this reason, flow data based on issued residence permits – despite it including valuable information on reason for migration and length of validity – is of limited use when focusing on intra-EU movements.



Reversed, this challenge in monitoring intra-EU flows also applies to EU nationals: being a EU citizen does not mean the starting point of person's most recent move to the Member State was from within the EU. By categorising all arrivals of EU nationals to a Member State as intra-European flows, we run the risk of overestimating their number. Still, when information on the previous country of residence is missing, our best guess is that the migrant is moving from their country of citizenship (or birth) (see Box 1).

Keeping the above limitations in mind, the total numbers of inflows and outflows are thoroughly reported on the Eurostat database dating back to 2008, with only a few exceptions (e.g. data for one or a few years missing for specific countries). Besides total numbers, flows reported by Eurostat generally include disaggregations by the country of birth and the nationality of immigrants, with some exceptions (Germany, Greece, France, Cyprus, Malta, Poland and Portugal). Aggregated values for EU-28 countries are available starting 2013; before that, it is difficult to aggregate values even manually because of missing country-level data for multiple countries.

Following EU Regulation, flow data should be possible to disaggregate by country of previous or (for emigration) next residence in Eurostat, but in practice this data is still missing for a number of countries.⁴⁴ The lack of this data severely limits researchers from tracking secondary movements within Europe (and tracking the exact number of de facto intra-regional movements). Furthermore, even in cases where the country of previous residence is available, Eurostat does not offer a further disaggregation of this data by citizenship or country of birth, limiting the complexity of information on the migration corridors highlighted. A double disaggregation by nationality and country of birth simultaneously is available if one of the two is categorised in a 'broad group' (such as 'EU28 countries except reporting countries'); as explained above and in Box 1, having at least this kind of option between nationality and country of previous residence would be extremely useful.

(continued on p. 35)

⁴⁴ CZ, DE, EL, FR, CY, LV, LU, HU, MT, PL, PT, RO



Box 1. Different definitions used in migration statistics, and the groups they delineate – a.k.a. why the availability of double/triple disaggregation by different definitions in migration flow statistics matters

The Venn diagram below illustrates how the different groups implicated in EU migration statistics relate to each other, including overlaps. Each of the three main definitions (black text) implicates four groups of migrants, depending on how they overlap with the other two definitions. The categories in white text indicate mutually exclusive groups.

Strictly speaking, statistics on ‘intra-EU movers’ should include the groups encompassed by the yellow circle: people moving from one EU country to another, including both EU nationals and non, born in the EU and not (II., III., V., VI.). When we simply look at statistics of immigrants to an EU country who are EU citizens, we risk including EU nationals who are in fact moving from a third-country at that point (I.), and excluding non-EU citizens who are moving from another EU country (V., VI.); the same argument applies to the EU-born definition. However, the size of the different subgroups (i.e., how non-overlapping areas and different overlaps compare to each other) can only be assessed if we can disaggregate the same data by two (ideally, three) definitions at the same time.

Examples for each group

Imagine an immigrant to the Netherlands (NL) who:

- (I.) Holds Italian citizenship but is in fact moving from Argentina, the country in which she was born
- (II.) A person born in Turkey who has since acquired German citizenship, but is now leaving Germany for NL
- (III.) A native Italian citizen moving from Italy to the Netherlands
- (IV.) A native Italian citizen who has been living in the United States, but now decides to move to NL
- (V.) A native Mexican citizen who has been living in Spain but now decides to move to NL
- (VI.) A son of Iraqi immigrants who was born in Hungary but has not been naturalized, moving from HU to NL
- (VII.) A son of Iraqi immigrants who was born in Hungary, has since moved to Iraq, but now moves to NL



(continued from p. 33)

In the case of years/countries for which the above Eurostat data is missing, UNDESA's database on migrant flows can be a highly practical supplement to Eurostat flow data as it includes some data for all 28 Member States, usually by both previous residence and citizenship. Nevertheless, the datasets have some drawbacks. Firstly, some country-level disaggregations are missing for some countries and years, but these do not completely overlap with Eurostat gaps, making it a useful complementary source. A second weakness is the format of the data, which is presented as a separate (long) Excel-table for every reporting country. The separate files are a setback for cross-country analysis, and the long format of tables require manual sorting-through to filter cases relevant to intra-EU movements. On a positive note, data are mostly available on a yearly basis for a long period of time (typically starting in the 1990s, but even 1980s in some cases), both by individual origin countries and aggregated by region, although it is important to note that the definition of 'Europe' includes some non-EU countries. Finally, double disaggregation by former residence and citizenship is not available (similarly to Eurostat).

For OECD countries – note that this excludes 6 Member States – the OECD Migration database can also be a good complementary source, both for total values and for disaggregation of flows by nationality. It is important to keep in mind, however, that the dataset does not track inflows of nationals, only foreigners. In some cases, individual countries' national statistical institutes (NSIs) also provide good complementary – or even overall better – statistics (e.g. Germany, United Kingdom – see Section III for more details).

Another noteworthy source for intra-EU migration flows is Abel and Sander's 2014 paper and resulting interactive website,⁴⁵ an impressive effort to estimate global migration flows between and within regions for five-year periods, 1990 to 2010, using UN stock data.⁴⁶

⁴⁵ <http://www.global-migration.info>

⁴⁶ G. J. Abel and N. Sander, 'Quantifying Global International Migration Flows', *Science* 343, no. 6178 (28 March 2014): 1520–22, <https://doi.org/10.1126/science.1251285>.



A disaggregation of flows by age and sex is typically available in Eurostat data. A breakdown by sex is available for all countries and years for which total flow figures are available. The availability of disaggregation by age, however, varies: it is entirely missing for Austria, Greece, Ireland, Romania, Slovenia and the UK and available since 2013 for Croatia, Lithuania, Luxembourg and Slovakia. In some cases – e.g. Austria, Ireland; see Section III – individual NSIs provide this data.

Further useful variables would include, for example, reason for migration, education and occupation background of incoming (and outgoing) individuals. Reason for migration is typically recorded in residence permits, but as mentioned above, this source of data is not applicable for EU citizens. According to our review, the one major EU-wide dataset that does contain this information is the Labour Force Survey's (LFS) 2014 (and 2008) ad-hoc module on labour migration.⁴⁷ Although LFS provides stock data, it includes years of residence and the information can therefore be transformed into flow data for the calculated years of arrival.

Information on the education and labour market participation of migrants is available both in the core module of LFS and Eurostat, but this data is generally collected to measure integration and thus refers to these characteristics at the time of the data collection, not at the time of migration. The same applies to a number of national databases as well, e.g. French and German NSIs record level of qualification and occupation, but again, at the time of data collection. Since these characteristics often change during the time spent in the host country, they cannot be accurately traced back using stock data and time of arrival.

One indirect way to calculate the education level of incoming (past) migrants – but this might only be applicable for a small number of observations – could be to use the variable *Year when the highest education was obtained* in the core module of LFS, which, compared to the years of residence for migrant respondents, can reveal if the education was obtained before or after moving to the host country – and in the former case, was likely the level of

⁴⁷ Data from the 2009 Eurobarometer and the 2004 EIMSS survey also includes reason for migration, but this data is outdated and/or not representative at the EU level. National surveys NELLS (Netherlands) and SOEP (Germany) also collect this information.



education at the time of migration. The one national-level survey that we found that includes information on migrants' usual occupation prior to migration is the UK's International Passenger Survey (IPS). The inclusion of questions regarding education and labour characteristics prior to migration in future migration-related EU-wide surveys would be a key step to identify how bilateral mobility patterns in Europe differ by skills, occupational status, and self-employment status.

B. Stocks of intra-EU migrants

Similarly to migrant flows, stocks of people who migrated from within the EU are reported alongside migrants arriving from third-countries in Eurostat's database (via NSIs, based largely on administrative records). Stocks of foreign nationals and population born abroad may be disaggregated by age and sex, but disaggregation by nationality and country of birth simultaneously is only available if one of the two is categorised in a 'broad group' (such as 'EU28 countries except reporting countries').

Statistics on the numbers of valid residence permits (and long-term residents) at a point in time also provide an additional insight into stocks of TCNs residing in Member States, in with additional information on the reason of migration and the length of stay. When focusing on stocks resulting from intra-European movements, however, we again must keep in mind that this refers not only to – and not to all – movements of EU nationals but rather EU nationals and TCNs' migrations from one Member State to another. In the case of stock data, this challenge is made more daunting by the fact that the countries of previous residence are not reported in Eurostat population data; Regulation 862/2007 requires Member States to record countries of previous residence for flow but not stock statistics.

Census 2011 (and following) data should be more useful in this regard: Regulation 763/2008 on population and housing censuses, requires Member States' census data to include inhabitants' previous place of usual residence and date of arrival in the current place; or place of usual residence one year prior to the census. Without exploring exactly what data was collected by each individual country, we rely on the ESS' Census Hub database (the



platform built do collect and distribute Census 2011 data across Europe), which includes variables for respondents' years of arrival and for their "residence one year before". The latter, however, only distinguishes internal and international movements (or no movements) and therefore cannot be used to determine the previous country of residence. Furthermore, the usefulness of the platform is greatly hindered by the limited options of variables that can be chosen jointly. For this reason, in our research we found it easier to download the data through the standard Eurostat data explorer platform.⁴⁸ Including information identifying the countries of previous residence in the census (or any other collection of other migration stock data) would help researchers access a key tool to comprehensively monitor intra-European movements of both EU nationals and TCNs.

As in the case of flow data – although to a lesser degree – a complementary source for Eurostat migrant stock data is provided by UNDESA's database on Trends in International Migrant Stocks. The database provides global bi-lateral stocks every 5 years between 1990 and 2015, disaggregated by sex. Most of the data are obtained from population censuses. The database has a global scope, but EU countries are easy to filter among receiving countries, which are listed by region. Origin of immigrants is defined as either country of birth (predominantly) or citizenship, but the two definitions are not distinguished for the figures in the dataset, and there are discrepancies regarding the inclusion of refugees; the inconsistency of definitions, in particular, significantly harms the precision of this data as a source for European migrant stocks.

Another complementary source is found in OECD's International Migration Database. This data faces the same geographical limitations as mentioned in the case of flows (6 EU countries are excluded). For the 22 countries that are included, however, OECD reliably reports stocks of foreign-born and foreign population – distinguishing between definitions, unlike UNDESA's stock data. Disaggregation by sex is also possible.

⁴⁸ Note that 'Population and housing census' category is separate from general population data; within this category, the Census 2011 sub-section has a separate 'folder' titled Population by migration characteristics.

Given these limitations of the current administrative data collections, an alternative way to identify intra-EU movers among migrant stocks is through surveys – keeping in mind the evident drawback of limited representativeness. Starting with the largest-scale survey, the core module of the LFS records the imperfect, but still useful variable *Country of residence one year before survey* (possible to combine with the *years of residence* variable for the current country); however, the limited information contained in LFS regarding country of nationality and country of birth should be kept in mind. EIMSS asked for all previous countries of residence; SOEP (Germany) also looks into past migrations in detail, as does wave 7 of the Understanding Society survey (UK).⁴⁹ REFLEX and HEGESCO asked for the country of residence at age 16. An added benefit of these surveys is that they also allow for disaggregation by various demographic and socio-economic characteristics. The obvious drawback is the lack of a migrant oversample (for the standard LFS, REFLEX, and HEGESCO) and the resulting very limited – and hardly representative – migrant stocks that can be identified through these surveys.

C. Statistics on selected aspects of intra-EU mobility

i. Reasons driving migration within the EU

Data on intra-European movers' reasons for migration can serve a key role in revealing the drivers behind the existence and prevalence of specific migration corridors within the EU; however, this information is currently not available at an EU level from administrative sources. Although the nature of migration is included – in good detail⁵⁰ – in residence permit records, as explained earlier, this data only refers to third-country nationals, and does not contain a tool for identifying intra-EU movers among them.

The most representative recent source of data on reasons for migration is the 2014 ad-hoc module of LFS, which asks this specifically. A close runner-up is the 2009 special wave (72.5) of the Eurobarometer, which works with a much smaller sample than LFS but has

⁴⁹ The UK's International Passenger Survey (IPS) also includes information on the previous country of residence, but given its nature the IPS provides flow data and is included in Eurostat's statistics, as it replaces national register data.

⁵⁰ Including not only reason for the immigrant receiving the permit (e.g. student, work, asylum), but additional breakdowns within these main reason (such as sector of employment).



questions on both motivations and disincentives to work abroad, as well as related attitude questions. An important difference is that most of the incentives and disincentives targeted by Eurobarometer 72.5 refer to a hypothetical future migration, not a recent migration. Nonetheless, the most recent past migration is also captured, including its location and questions about the type of work performed during that migration. Additionally, interviewees are asked if they have ever studied abroad, and commuted across borders (but these destinations are not specified). A weakness of both the LFS and the Eurobarometer is that only have approximate ways to capture the country of previous residence (LFS: broad groups for country of citizenship, birth, and residence 1 year prior, as well as last country of work abroad; Eurobarometer: citizenship, destination of last move, duration of that move). Therefore, they are imperfect for capturing the drivers of intra-EU mobility specifically.

The EIMSS survey was well-designed for these purposes in that identified not only the reason for settling in the destination country, but also the country of residence prior to migration – but, again, this data is only available for 2004 and for a handful of countries.

Among national statistical institutes, Italy, Greece and Netherlands record the reason for migration in their databases. Importantly, these databases are available online (except for the Netherlands, where previous authorization is necessary) and are downloadable in English. The reason for migration is also available, albeit in census data only, in Croatia and Slovakia. Among national surveys, SOEP (Germany), the UK Understanding Society survey's wave 7 (2014), and UK International Passenger Survey contain information on both the reason for migration and the country of previous residence, while NELLS (Netherlands) only includes the former.

Labour migration

Focusing on European labour migration in particular, the most useful EU-level database to consult is undoubtedly the Labour Force Survey. The core questionnaire allows to identify foreign citizens in the surveyed country while also asking detailed questions about their labour characteristics; while this is not strictly speaking labour migration in that it is not specified if labour was the motivation behind the move, it does give us some information on



EU citizens working in other EU countries (keeping in mind the limited information on their countries of origin). The questionnaire also asks for the country of place of work, which, if different than country of residence, allows to identify cross-country commuters. The 2008 and the 2014 ad-hoc modules are particularly relevant to assess the prominence of labour-motivated migration (together with the socio-demographic profile of these movers), since they ask about *Reason for migration* and *Last country of work abroad*.

The Eurobarometer's 2009 special wave on *Geographical and labour market mobility* also contains useful information: it asks if the respondent she has ever lived or worked abroad, and if yes (including if they are currently doing that); the location and duration of that migration; the type of work she has done while abroad (if any); how she found that job; a variety of similar but also even more detailed questions regarding an intended future move for work purposes. As shown by these questions, this special wave has addressed some key questions related to labour mobility within the EU which were a highly useful supplement to the data provided by the LFS ad-hoc module (in fact, one might argue that it went into more depth than the LFS did). Repeating this special wave in the near future, in a post-recession context, would no doubt yield valuable data, in particular to understand current EU labour migration attitudes and intentions.

In terms of country-specific databases on EU labour mobility, an example that stands out is Germany, specifically the Research Data Centre (FDZ) of the German Federal Employment Agency (BA), which has a variety of microdata originating from administrative registers and surveys with vast information on migration and labour market history variables.

Student migration

Statistics on tertiary student mobility within Europe are available through two main sources: for long-term (degree) mobility, the joint data collection effort of education data by UNESCO, OECD and Eurostat (UOE); for short-term mobility (one to two semesters), the database of the Erasmus student exchange programme.



Starting with UOE, the results regarding Europe (EU and EEA/EFTA) are shared on the Eurostat database.⁵¹ Following OECD practice, this collection of data distinguishes between resident ‘foreign students’ (who are residents in the country as a result of a prior migration) and non-resident, foreign ‘mobile students’, who came to the country explicitly to pursue an education programme.⁵² This helps capture the real volume of education-driven mobility without confusing it with the enrolment statistics of non-naturalised migrant residents.

Until the reference year 2012, data is presented in separate datasets for mobile and foreign students. For both groups, disaggregation is available within country-level data by enrolled students’ level of education and their sex, field of education, origin, or destination (for emigrants). Note that the separation of these breakdowns does not allow researchers to disaggregate specific country-to-country flows by characteristics other than level of education. Mobile students’ data includes a dataset on graduates, but only by level of education and sex, not origin – which prevents the identification of migration corridor.

The data from 2013 onwards only refers to mobile students and graduates. For both groups, migration corridors (as defined by country of analysis and country of origin) can be broken down by education level and sex. This is an improvement on the earlier system, but it would still be useful to include a dataset that allows for disaggregation both by country of origin and field of education.⁵³ The UOE datasets are therefore a reliable source of data on intra-

⁵¹ The basis for this data collection is set in Regulation (EC) No 452/2008. Data was collected based on a gentlemen’s agreement until 2011, Regulation (EU) No 88/2011 for 2011 and 2012 data, and Regulation (EU) No 912/2013 from 2013 onwards.

⁵² OECD, ‘Foreign / International Students Enrolled – Metadata’, OECD.Stat, accessed 29 June 2017, <https://stats.oecd.org/Index.aspx?DataSetCode=RFOREIGN>.

⁵³ The country of origin in the learning mobility data should, in principle, refer to the country of prior secondary education. Until the 2015 reference year, countries were allowed to use country of prior residence or citizenship or another concept instead of country of prior education. From the 2016 reference year onwards all countries are supposed report data according to the country of prior education [Eurostat, ‘Education Administrative Data from 2013 Onwards (ISCED 2011)’, 2017, http://ec.europa.eu/eurostat/cache/metadata/en/educ_uoe_enr_esms.htm.]



EU mobility of university students (as well as mobility into and from the EU), covering all EU-28 countries with some exceptions.⁵⁴

Moving on to Erasmus data, mobility statistics starting from the academic year 2008-2009 are available through the European Union Open Data Portal website. The shared data includes not only study exchanges but also work placements as well as teaching assignments and staff training. Raw microdata is freely available to download and includes disaggregations by sending and receiving country, age, gender, grant, duration, subject area, level of study, and other characteristics. It is therefore possible to use this data to identify and describe migration corridors of short-term student mobility across Europe.

Survey data from PISA and HEGESCO/REFLEX contain some additional information that might be useful for studying intra-European education mobility. As discussed in Section II., PISA provides triennial stock data – although through a limited sample – on 15-year-old students born abroad, including how long they have been residing in the country. HEGESCO and REFLEX, on the other hand, provide data on the migration trajectories of former higher education graduates (5 years on), including their country of birth, country of residence at age 16, during higher education studies, when first starting employment, and at the time of the surveys. The limitations of these datasets are that they are over a decade old and their migrant subsample is likely to be very small.

Finally, as mentioned above, the 72.5 wave of Eurobarometer (2009) asks respondents if they have ever studied abroad, but without asking to specify the location if yes (nor an indication of time). Therefore, while these are retrospective education mobility-related statistics about respondents who are certainly European residents, we cannot know if the indicated mobility was intra-European.

⁵⁴ Note: data for mobile students (by country of origin, 2013-2015) is missing for Greece; data for degree mobile graduates is missing for Greece, Spain, France, and Poland. Data for mobile students (by country of origin, 2008-2012) is missing completely for Czech Republic, Greece, France, Italy.



iii. Irregular migration

Flows and stocks of undocumented migrants, by definition, tend to remain undetected in administrative records; even when data collection is designed in a way that allows for the inclusion of unregistered inhabitants (e.g. censuses, non-register based surveys), their irregular migrant status is unlikely to be recorded unless specifically asked. Major collections of migrant stock and flow statistics such as Eurostat or OECD tend to exclude irregular migrants altogether from their enumerations, as do surveys that draw their sample using registers (e.g. SOEP, LFS in some countries); meanwhile, censuses and non-register based surveys (such as PISA, or LFS in other countries) may include respondents with an irregular status, but they remain unidentified as such in the data.

It is important to note that the concept of irregular migration within the EU only really applies to TCNs: while EU/EEA nationals are usually required to register their residence when moving to another country for more than three months, failing to do so may subject them to a fine but not to be expelled.⁵⁵ In fact, irregular migrants are defined as “third-country nationals who do not fulfil, or no longer fulfil, the conditions of entry as set out in Article 5 of the Schengen Borders Code or other conditions for entry, stay or residence in that Member State”.⁵⁶

There are multiple ways for third-country nationals to enter irregularity even if they fulfilled regulations when they first arrived to the EU. Examples include staying after their visa or residence permit has expired, becoming employed without a work permit, entering some non-Schengen countries without an EEA entry visa (only holding a visa/residence permit from a Schengen country) etc.⁵⁷ Non-EU family members of EU nationals are often subject to different regulations, as are citizens of certain non-EU countries, which further complicates the task of monitoring irregularity of flows within Europe.

⁵⁵ Europa.eu, ‘Registering Your Residence Abroad’, Europa.eu - Your Europe - Citizens, 2017, http://europa.eu/youreurope/citizens/residence/documents-formalities/registering-residence/index_en.htm.

⁵⁶ Giulio Sabbati and Eva-Maria Poptcheva, ‘Irregular Immigration in the EU: Facts and Figures’, *Briefing*, European Parliamentary Research Service, 2015, https://emnbelgium.be/sites/default/files/publications/eprs_briefing_0.pdf para. 1.

⁵⁷ For a detailed explanation, see Jandl, Vogel & Iglicka ‘Report on Methodological Issues’, *Research Paper*, CLANDESTINO Undocumented Migration: Counting the Uncountable, 2008., p.11



Therefore, the standardised EU data collection on irregular migration statistics focuses solely on non-EU citizens. Categorised under *Enforcement of immigrant legislation*, Eurostat provides statistics on TCNs who were refused entry at the external borders of the EU, found to be illegally present in the Member State's territory, were subject to an obligation to leave the territory of Member States, or have left the Member State's territory to a third country.

Law enforcement statistics, however, only capture a fraction of the irregular immigrant population – namely, those who were “caught”. The CLANDESTINO Database on Irregular Migration aimed to bridge this gap by combining enforcement data with regularisation data, support service data, administrative data, expert, migrant, and employer surveys and other sources. This CLANDESTINO the most comprehensive existing database on irregular migration in Europe, to our knowledge. The database contains data and estimates from 12 EU countries⁵⁸ mostly between 2007 to 2009, but information is still uploaded occasionally. Besides figures and analysis, the project includes a quality classification assessing the reliability of each of its estimates.

In sum, the statistics available through Eurostat and CLANDESTINO are useful indicators of flows into and out of, as well as stocks inside the territory of the EU of TCNs with an irregular status. However – even though they may include such cases – these statistics do not provide specific information on the intra-EU movements of migrants with an irregular status. Such information remains a data gap to be addressed.

iv. Monitoring lifetime/multiple migrations

Tracking lifetime migration trajectories – or even just multiple migrations – of EU residents with existing data is a major challenge we encountered in our mapping exercise.

As discussed with regards to migrant stocks, it is often difficult to know even the migrant's most recent previous country of residence, the only information on a country of origin being the migrant's country of birth or citizenship. This is the case for migration data based on population stock statistics, residence permits, and law enforcement data (as presented in the Eurostat database), as well as OECD stock and flow data and the PISA survey, to mention

⁵⁸ AT, CZ, FR, DE, EL, HU, IT, NL, PL, SK, ES, UK



a few. The German NSI represents a similar but slightly more informative practice in that it includes information on migrant status, age at the time of entry, duration of residence, citizenship and naturalization status, as well as former citizenship.

A limited number of sources include the migrant's country of previous (or next) residence, revealing the exact trajectory of one migration, at least. However, this type of source contains no information on other past migrations. In principle, this is the case for Eurostat's flow statistics, although in reality this data is still missing for a significant number of countries. In a somewhat similar method, the 2014 ad-hoc module of LFS asks for the last country of work abroad (within the last 10 years), but without a question to specify when that migration took place, or if other migrations not involving employment have taken place since. The 2009 special wave of the Eurobarometer (72.5) asks for the most recent migration, including the location and the length of stay, but no information on how long ago that was compared to the time of the survey. An outstanding country-specific example is Italy's NSI, which records not only the previous place of usual residence, but also residence a year prior and residence five years prior. The UK (through its Passenger Survey) records country of last or next residence, along with previous main reason for migration.

Another group of datasets provides information on migration histories by inquiring the country of residence one year prior to the time of data collection. This is better than no information, but is far from ideal in that it only tracks one, very recent migration (for individuals or households residing in the destination country for no more than a year) and does not exclude the possibility that the migrant arrived to the present country from another country (which he or she moved to and away from within one year). Data on the country of residence one year prior is included in the core module of LFS. Furthermore, according to current regulations, EU censuses (i.e. Census 2011) are required to include data on either respondents' previous place of residence (and date of arrival) or place of usual residence on year prior to the census.⁵⁹ The former variable would clearly be more valuable to capture exact migration trajectories, but the Census Hub database only contains the

⁵⁹ Eurostat, 'Census 2011 Round (Cens_11r) Reference Metadata in Euro SDMX Metadata Structure (ESMS)'.



latter variable and does not differentiate between individual countries (as mentioned previously). Individual country censuses, if accessible, may therefore contain useful information, but – whether collected or not – the current overarching database excludes this data. Some NSIs' databases, such as France and Spain, also record residence a year prior to the time of data collection.

A similar method asks for the individual's country of residence at one point in the past, not one but multiple years prior, or, alternatively, when the respondent was a specific age.⁶⁰ Again, this type of data is useful in providing at least some information on the individual's migration history, but it does not allow the researcher the map specific country-to-country trajectories since it leaves out potential migrations that happened between the years for which residency was recorded in the data. The likelihood of the data accurately capturing all migrations is higher when places of residence are asked for multiple life stages which are directly connected (or almost). HEGESCO and REFLEX, for instance, provide fairly good data as they ask about country of birth, country of residence at age 16, during higher education studies, when first starting employment, and at the time of the surveys (5 years following graduation). Unfortunately, the migrant subsample can be expected to be very limited within these surveys as it did not specifically target mobile individuals.

An exceptional example among data sources tracking lifetime or multiple migrations is the aforementioned special wave of the Eurobarometer (72.5), which asked detailed questions about a planned *future* move,⁶¹ including questions assessing the maturity/certainty of these plans. This is valuable data in that it uniquely outlines future possible migrations while also containing information on past (and/or present) migratory behaviour of the individual.

In summary, the ideal system of data collection on lifetime migration is as thorough as possible; the best existing example we found in the European context is EIMSS, which tracks

⁶⁰ It is also worth noting the opposite method applied in the UK Passenger Survey, which instead of asking about the respondents' location at a specific time, asks to specify the year when he or she last moved to or away from the UK.

⁶¹ Note: Wave 3 and 7 of the UK Understanding Society survey also address planned future migrations, although the latter wave does specify a destination.



every other country the respondent has lived in, prior migration to the current destination country, reasons for settlement, and future moving aspirations including retirement. The German SOEP is also very well-designed for these purposes, as it asks respondents to specify every single move (including dates) since birth, including potential moves away from and then back to Germany. Another outstanding national survey in this regard is the UK's Understanding Society survey (wave 1 & 7), which investigates all prior migrations in both native and migrant respondents' lives (although it does not include all dates, only whether it was prior to or following any move to the UK). The implementation of such surveys at the broader European level – or the inclusion of prior migration questions to existing EU-wide questionnaires such as LFS – would provide crucial lacking information on extended migration trajectories and the prevalence of multiple migrations over European inhabitants' lifetimes.

v. Short-term migration, circular migration, cross-border commuting

Short-term migration – a change of residence to another country for a length of over 3 to under 12 months ⁶² – is particularly difficult to track with existing European statistics. Residence permits do include information allowing to identify short-term migrants among third-country nationals (and only them), but not to capture intra-EU movers among them. Passenger surveys seem to be the best available data source to observe the volume of short-term movements (as well as tourism) from specific European countries; unfortunately, to our knowledge, at this time only the UK and Cyprus carry out these types of surveys systematically.

Similarly, we struggle to find internationally available sources on circular migration within Europe. This challenge was thoroughly discussed in UNECE's recent report *Defining and Measuring Circular Migration*, which describes definitional issues as well as a detailed review of potential sources that may be exploited to build circular migration statistics, using

⁶² OECD, 'OECD Glossary of Statistical Terms - Short-Term Migrant Definition', 2003, <https://stats.oecd.org/glossary/detail.asp?ID=2457>.



Italy and Sweden as in-depth examples; we recommend referring to this report for a comprehensive overview on the topic.⁶³

The availability of cross-border commuting data is slightly better, as the core module of LFS includes a question on the country of place of work (which can be compared with country of residence). Additionally, the special wave (72.5) of the 2009 Eurobarometer, asks whether the respondent commutes to work in another country, the frequency of the commute, and further questions about hypothetical or intended international work commutes. A major shortfall of this data is that the destination of the commute is not identified. Furthermore, both for the LFS and the Eurobarometer, we may expect a small sample as cross-borders commuters are not specifically targeted.

Before moving on to conclusions and recommendations, Table 1-2 on the next page summarises the main sources we identified in our review of available data, split by main and complementary sources. While the table may serve as a quick summary/overview of the data sources and limitations discussed in throughout this section it is by all means just a quick summary – for an in-depth analysis and full accuracy, please refer to the text above instead.

⁶³ UNECE, 'Defining and Measuring Circular Migration' (New York and Geneva: United Nations Economic Commission for Europe (UNECE), Task Force on Measuring Circular Migration, 2016).



Table 1. Main sources for statistics on selected aspects of intra-EU migration (1/2)

	Flows	Stocks	Reason for migration	Labour migration	Student migration
Main source	Eurostat	Eurostat ⁴	LFS 2014 AHM ⁴	LFS ^{4,5}	UOE
Complementary sources	UNDESA	OECD ^{2,4}	Eurobarometer 72.5 ^{1,2,4}	LFS 2014 AHM ⁴	Erasmus data
	OECD ^{2,4}	Census 2011 ^{1,4}	EIMSS ^{1,2}	Eurobarometer 72.5 ^{1,2,4}	PISA ⁵
	Abel & Sander (2014)	UNDESA ⁴	Eurostat (Residence permits) ⁴		HEGESCO/REFLEX ^{1,2,5}
		LFS ^{4,5} , LFS 2014 AHM ⁴	Understanding Society (w7) ³ UK IPS ³		
		EIMSS ^{1,2}	SOEP ³		
			NSI data: ³ IT, EL, NL	NSI data: ³ DE	

Table 2. Main sources for statistics on selected aspects of intra-EU migration (2/2)

	Irregular migration	Lifetime/ multiple migrations	Short-term migration	Cross-border commuting
Main source	-	-	-	LFS ^{4,5}
Complementary sources	Clandestino ^{1,2,4}	LFS ^{4,5}	Eurostat (Residence permits) ⁴	Eurobarometer 72.5 ^{1,2,4}
	Eurostat ⁴	Eurobarometer 72.5 ^{1,2,4}	Erasmus data	
		EIMSS ^{1,2}	UK IPS ³	
		SOEP ³		
		Understanding Society (w1,3,7) ³ UK IPS ³		
		NSI data: ³ DE, FR, ES		

Limitations of data sources indicated in superscripts, numbers indicating:

1. Old data
2. Not all EU-28
3. National only
4. Problem identifying intra-EU movements
5. Lack of migrant oversample



Conclusion

The aim of this paper was to provide a discussion and overview of the main databases available to understand mobility within the European Union. Following some key questions of interest, we mapped existing data sources and evaluated their usefulness and quality in supporting intra-EU migration research.

Our initial overview of the different types of data sources (e.g., administrative registers versus surveys) was greatly aided by previous comprehensive reviews borne by projects such as THESIM and PROMINSTAT.^{64,65} Guided by the criteria outlined by these works as well as our own research priorities and experiences, we identified the best – most comprehensive and practical in use – databases, highlighting the main data sources for each aspect of EU mobility, followed by some other sources that can best complement the gaps left by the main data (no data source we found was exhaustive). The Eurostat database, the Labour Force Survey (both core and 2008/2014 ad-hoc modules), the migration databases of UNDESA and OECD, as well as the EIMSS survey and the special wave 72.5 of the Eurobarometer were the among the most useful sources we identified. Besides presenting these main sources, however, an equally critical objective of our exercise was to shed light on the remaining challenges and limitations of measuring intra-EU mobility. Below we recapitulate our main conclusions and recommendations.

In summary, we find that the primary challenge in finding data on intra-EU migration relates to identifying the part of EU-related mobility that truly takes part within the external borders of the EU. For recent years, data on the citizenship and/or country of birth of migrants is usually available, but not on the country of residence prior to migration; most data sources are either not interested in the country of departure seem to be willing to assume that it is the same one as the country of citizenship or birth. This makes it very difficult to accurately identify migration corridors within Europe – both for EU nationals and third-country citizens. Higher rates of compliance with existing EU requirements to collect

⁶⁴ Poulain, Perrin, and Singleton, *THESIM*.

⁶⁵ Kraler and Reichel, *Statistics on Migration, Integration and Discrimination in Europe. PROMINSTAT Final Report*.



information on residents' previous/next country of residence for migration flow statistics on behalf of Member States would be a major step in filling this gap. Additionally, the inclusion of a question regarding country of previous residence in either the LFS core module or at least its upcoming ad-hoc module on labour migrants would be key to gain this information as it relates to existing migrant stocks.

Next, a significant obstacle to in-depth research being conducted with Eurostat data is the separation of available disaggregations. While it is understandable that microdata is not available on this platform, many useful breakdowns of the data by important characteristics (e.g. age, sex, country of birth *and* citizenship) are already collected but often displayed separately, preventing multiple disaggregations which could be key for analyses. Keeping in mind the data limits of online interactive tables, it should be possible to offer datasets including all available disaggregations for a given statistic in the online bulk download facility of Eurostat.

Another challenge to in-depth studies on intra-EU mobility is the lack of information on lifetime migration (or at least multiple migrations) of intra-EU movers. This information could probably be best gathered by repeating and extending the EIMSS survey (conducted in 2004), which covered all previous migrations of respondents. A repetition of the 72.5 (2009) special wave of the Eurobarometer would also undoubtedly yield some valuable data in this regard.

Further themes on which there is limited or no data available in the intra-EU context include irregular migration, short-term migration, cross-border commuting and circular migration. In addition to these, acquiring more background information on migrants, especially data on their skills, occupation and other socio-economic characteristics at the time of migration could help researchers gain a better understanding of which EU residents are moving to which Member State; this could in turn help policymakers understand drivers and anticipate needs of new inhabitants.



References

- Abel, G. J., and N. Sander. 'Quantifying Global International Migration Flows'. *Science* 343, no. 6178 (28 March 2014): 1520–22. <https://doi.org/10.1126/science.1251285>.
- Council of the European Union, and European Parliament. '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 (Text with EEA Relevance)'. *Official Journal of the European Union*, no. L 199/23 (31 July 2007): 23–29.
- . 'Regulation (EU) No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European Demographic Statistics Text with EEA Relevance'. *Official Journal of the European Union*, no. L 330/39 (12 October 2013). <http://data.europa.eu/eli/reg/2013/1260/oj>.
- Eurofound. 'Analysis of the Socioeconomic Situation of Migrants – Gathering Comparable Data on Their Living Conditions'. Workshop report. Brussels: Eurofound, 17 November 2010. http://csdle.lex.unict.it/Archive/LW/Data%20reports%20and%20studies/Reports%20and%20studies%20from%20EUROFOUND/20110325-115319_Eurofound_Migrants_March11pdf.pdf.
- Europa.eu. 'Registering Your Residence Abroad'. Europa.eu - Your Europe - Citizens, 2017. http://europa.eu/youreurope/citizens/residence/documents-formalities/registering-residence/index_en.htm.
- European Commission. 'Commission Implementing Regulation (EU) No 205/2014 of 4 March 2014 Laying down Uniformed Conditions for the Implementation of Regulation (EU) No 1260/2013 of the European Parliament and the Council on European Demographic Statistics, as Regards Breakdowns of Data, Deadlines and Data Revisions Text with EEA Relevance'. *Official Journal of the European Union*, no. L 65/10 (3 May 2014): 10–26.
- . 'Geographical and Labour Market Mobility'. Special Eurobarometer 337. European Commission, June 2010. <http://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/803>.
- . 'Report from the Commission to the European Parliament and the Council on the Implementation of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection'. Brussels: European Commission, 20 September 2012.
- . 'Report from the Commission to the European Parliament and the Council on the Implementation of Regulation (EC) No 862/2007 on Community Statistics on Migration and International Protection'. Brussels: European Commission, 30 July 2015.
- European Social Survey. 'Attitudes towards Immigration and Their Antecedents - Question Design Final Module in Template'. Centre for Comparative Social Surveys, City University London, 2015.



- http://www.europeansocialsurvey.org/docs/round7/questionnaire/ESS7_immigration_final_module_template.pdf.
- Eurostat. 'Census 2011 Round (Cens_11r) Reference Metadata in Euro SDMX Metadata Structure (ESMS)', 2014.
http://ec.europa.eu/eurostat/cache/metadata/en/cens_11r_esmscs.htm.
- . '"Dublin" Statistics (Migr_dub) Reference Metadata in Euro SDMX Metadata Structure (ESMS)', 2010.
http://ec.europa.eu/eurostat/cache/metadata/en/migr_dub_esms.htm.
- . 'Education Administrative Data from 2013 Onwards (ISCED 2011)', 2017.
http://ec.europa.eu/eurostat/cache/metadata/en/educ_uoe_enr_esms.htm.
- . 'EU Statistics on Income and Living Conditions (EU-SILC) Methodology – Sampling'. Eurostat Statistics Explained, 2016. [http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_\(EU-SILC\)_methodology_-_sampling#Sample_size](http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_sampling#Sample_size).
- . 'European Union Labour Force Survey: Description of the Dataset', 2017.
<http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey>.
- . 'International Migration Statistics Reference Metadata in Euro SDMX Metadata Structure (ESMS)', 2017.
http://ec.europa.eu/eurostat/cache/metadata/en/migr_immi_esms.htm.
- Jandl, Michael, Dita Vogel, and Krystyna Iglicka. 'Report on Methodological Issues'. *Research Paper, CLANDESTINO Undocumented Migration: Counting the Uncountable*, 2008.
- Kraler, A., and Reichel. *Statistics on Migration, Integration and Discrimination in Europe. PROMINSTAT Final Report*. Brussels: European Commission, 2010.
- Lien, H., and I. Toleikyte. 'The Labour Market Situation of Migrants and Their Immediate Descendants Evaluation of the 2014 Labour Force Survey Ad Hoc Module'. Luxembourg: Eurostat, December 2015.
http://ec.europa.eu/eurostat/documents/1978984/6037334/Evaluation_report_AHM_2014.pdf.
- OECD. 'Foreign / International Students Enrolled – Metadata'. OECD.Stat. Accessed 29 June 2017. <https://stats.oecd.org/Index.aspx?DataSetCode=RFOREIGN>.
- . 'OECD Glossary of Statistical Terms - Short-Term Migrant Definition', 2003.
<https://stats.oecd.org/glossary/detail.asp?ID=2457>.
- . 'Statistical Annex'. In *International Migration Outlook 2016*. OECD Publishing, 2016.
http://dx.doi.org/10.1787/migr_outlook-2016-46-en.
- Poulain, Michel, Nicolas Perrin, and Ann Singleton, eds. *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: Presses universitaires de Louvain, 2006.
- Recchi, Ettore, Emiliana Baldoni, Francesca Francavilla, and Letizia Mencarini. 'Geographical and Job Mobility in the EU (Final Report)'. 'Empirical Evidence on Job and Geographical Mobility in the European Union' Tender No. VT/2005/0107 DG Employment, Social Affairs and Equal Opportunities. CIUSPO, University of Florence, June 2006.



- Sabbati, Giulio, and Eva-Maria Poptcheva. 'Irregular Immigration in the EU: Facts and Figures'. *Briefing. European Parliamentary Research Service*, 2015.
https://emnbelgium.be/sites/default/files/publications/eprs_briefing_0.pdf.
- UNECE. 'Defining and Measuring Circular Migration'. New York and Geneva: United Nations Economic Commission for Europe (UNECE), Task Force on Measuring Circular Migration, 2016.





REMINDER

ROLE OF EUROPEAN MOBILITY AND ITS IMPACTS
IN NARRATIVES, DEBATES AND EU REFORMS

The REMINDER project is exploring the economic, social, institutional and policy factors that have shaped the impacts of free movement in the EU and public debates about it.

The project is coordinated from COMPAS and includes participation from 14 consortium partners in 9 countries accross Europe.



This project has received funding from the European Union's Horizon 2020 research & innovation programme under grant agreement no 727072

