**Looking for immigrants in the European Labour Force Survey and the EU census: a comparison based on the 2011 figures**

**Keywords:** foreign-born population, EU Census, EU LFS, data comparison and validation

# Introduction

Labour mobility refers to the extent the workers are able or willing to move between different jobs, occupations, and geographical areas. In the specific context of the European Union (EU), as the individuals have the right to freedom of movement between countries, it refers mainly to inter-country geographical mobility. This has driven important labour force movements across the EU territory. As such, the active participation of migrants and of their immediate descendants in the EU labour market and, generally speaking, in daily life is very important for social cohesion and economic efficiency.

An increase in labour mobility in the EU, coming both from other Member states and from countries outside its territory, creates a need for detailed and up-to-date statistical information to feed current political discussions on labour market policies in general and on labour mobility in particular. The two main data sources for obtaining information on the topic are the EU Labour Force Survey and the Population and Housing Census, supplemented by yearly administrative population data.

# Methods

**2.1 Objective and method**

The objective of this paper is to compare the data obtained from the EU-LFS and the EU Census regarding characteristics of the foreign-born population residing in the EU. Combining these two data sources can bring new opportunities for analysing this group. The purpose is therefore to provide a clear overview of the differences and similarities when comparing the results coming from the two data sources for the same reference year, 2011. Potential explanations for the differences identified are also put forward. The main question is whether or not it would be possible to update the more structural information in the EU Census with the more timely data available in the EU-LFS when publishing statistics about the foreign born population.

**2.2. Main data sources**

The main data source for measuring labour mobility is the European Union’s Labour Force Survey (EU-LFS), the largest EU sample survey covering the resident population aged 15 and over, living in private households and residing in the EU. It provides detailed quarterly and annual data on employment and unemployment, broken down along many dimensions. One of the major benefits of having a large sample size is to be able to study the characteristics on the labour market of the foreign-born population (i.e. persons born in a country other than her/his country of residence and whose residence period in the host country is, or is expected to be, at least 12 months). The current situation is still in need of improvement as, even if the sample size is large, the desired level of detail cannot be achieved all the time due to a combination of several factors: the sample was not specifically designed to follow immigrants, the migration phenomenon is still relatively marginal in absolute numbers (e.g. in 4 countries this population counts for less than 30 000 persons), the patterns (origin countries versus destination countries) and size largely vary across EU Members States (e.g. from about 11 000 persons to 11 000 000 persons). In addition to the fact that the sample is not specifically designed to capture well the immigrants, there can also be significant delays for their presence in sampling frames due to mainly time delay necessary for registering/deregistering.

Population and housing census bring the powerful strength of an exhaustive survey and provide the most reliable and geographically detailed count of the population and cross-classify it in great detail for a set of selected characteristics.

The EU census data are disseminated in an aggregated form which translates into a set of multidimensional cross-tabulations called "hypercube". Their structure are defined in the legal implementing rules for the EU census and agreed with all Member States. More precisely, they are result of the extensive planning and close cooperation and consultations between Eurostat and EU Member States. Census Hub – the data collection and dissemination platform – differs substantially from traditional Eurostat on-line dissemination platform, the main differences being that the data are (a) physically stored in Member States, (b) owned by Member States and published under their own responsibility. In this context, it is important to note that the EU Census's output is restricted to the well-defined set of hypercubes, no customised data extraction being possible as it is the case for EU-LFS, where Eurostat has access to microdata.

# Results

First, the conceptual consistency is analysed. The conclusion is that while variables definitions and categories used ensure all premises for a good convergence of the two data sources, the differences in data collection particularities and data treatments can potentially introduce important differences in data when comparing EU-LFS and EU Census.

After this step is completed, the corresponding multidimensional cross-tabulations are extracted from the two data sources. The share of each category in total corresponding population (i.e. either total population or total foreign-born population) is computed and the figures referring to the same category based on the EU Census and EU-LFS are compared. In order to assess the similarity of the two data sources, we look at the difference between them in percent points by subtracting category's share according to EU-LFS estimates from the corresponding share according to EU Census data. For instance, for a given category of a dimension, the percentage point difference is computed as:

$PERCENT DIFFERENCE\_{country i}=SHARE\_{country i}^{Census}-SHARE\_{country i}^{EU-LFS}$

As a rule of thumb, we consider that the two data sources are very similar (dark green) if the difference does not exceed ±2 pp, similar (light green) if the difference is more than ±2 pp but less than ±5 pp and not similar (red) if it exceeds ±5 pp. This also takes into account the fact the EU-LFS is a survey; therefore the point estimates produced are within a confidence interval.

The dimensions used for cross-tabulations are reporting country and migration status on one hand and a third variable on the other hand (age or sex or attainment education level or working status). For instance, the foreign-born population (migration status dimension) in a country is split by the three age groups. In addition to these cross-tabulations; a comparative table for active and employment rates is presented. All detailed cross-tabulations and their percent differences are presented in the paper.

Table 1.: Percent difference between the same categories identified in the EU-Census and the EU-LFS in the same reference year, 2011



# Conclusions

The current situation is not favourable to combining the two data sources for getting more frequently estimates for foreign-born population. Despite the high power of EU-LFS to capture the foreign-born population important weakness arise when this population is broken down by some of the dimensions analysed in this paper, and in particular those that are of high relevance for looking at their labour market situation.

## The recommendation is to look for solutions to substantially increase the power of capturing foreign-born population in a sample so that it can be analysed by its multiple facets, and to collaboratively look for best practices for better collecting this population that is of very high relevance in a policy perspective.