



# DEMOGRAPHIC RESEARCH

*A peer-reviewed, open-access journal of population sciences*

---

## ***DEMOGRAPHIC RESEARCH***

**VOLUME 42, ARTICLE 27, PAGES 763–776**

**PUBLISHED 30 APRIL 2020**

<http://www.demographic-research.org/Volumes/Vol42/27/>

DOI: 10.4054/DemRes.2020.42.27

*Research Material*

### **A visual tool to explore the composition of international migration flows in the EU countries, 1998–2015**

**Beata Nowok**

This publication is part of the Special Collection “Data Visualization,” organized by Guest Editors Tim Riffe, Sebastian Klüsener, and Nikola Sander.

© 2020 *Beata Nowok*.

*This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.*

*See <https://creativecommons.org/licenses/by/3.0/de/legalcode>.*

## Contents

1	Introduction	764
2	Data on international migration flows	764
3	Interactive ternary plots	765
4	Illustrative examples	767
5	Concluding remarks	773
	References	774

## **A visual tool to explore the composition of international migration flows in the EU countries, 1998–2015**

**Beata Nowok<sup>1</sup>**

### **Abstract**

#### **BACKGROUND**

Ternary plots can effectively display compositional data and help to understand its multidimensional patterns. However, they are currently greatly underused in demography.

#### **OBJECTIVE**

My goal is to develop an interactive web-based visualization of compositional data on international migration flows in order to compare its developments over time and across countries, and through this example demonstrate the utility of ternary diagrams.

#### **METHODS**

R Shiny framework is used to build a web application. Immigration and emigration flows in the EU Member States from 1998 to 2015 are grouped into three conventional categories (nationals of a reporting country, [other] EU nationals, and non-EU nationals) and presented on ternary diagrams.

#### **RESULTS**

Compositional data on migration flows can be effectively visualized using ternary plots. An interactive web application has been developed that allows comparative exploratory analysis of immigration and emigration composition and size in the EU countries over time. The impact of the entry of new countries to the EU can be assessed by comparing data referring to the EU composition of 28 Member States with data referring to the EU composition of the reference period.

#### **CONCLUSION**

Ternary plots can facilitate compositional analysis of migration flows grouped into three categories.

#### **CONTRIBUTION**

I introduce ternary plots for an exploratory analysis of migration composition and provide an online tool to carry out illustrative exploration. Demographic analysis of compositional data can benefit from the wider use of such plots.

---

<sup>1</sup> School of Geosciences, University of Edinburgh, UK. Email: [beata.nowok@ed.ac.uk](mailto:beata.nowok@ed.ac.uk).

## 1. Introduction

Population composition and its changes are central to demographic analysis. There are countless examples of visual display of proportions, but graphical solutions are limited. The most commonly used ones are proportional stacked bar plots, stacked area charts, and pie charts. Sometimes proportions are also presented using multiset bar charts or line charts, but then the sense of a whole and its parts is lost. A recent example by Schöley and Willekens (2017) of encoding compositional data on the Lexis surface demonstrates the usefulness of more elaborate visualization techniques as a platform for data exploration. My aim is to compare compositions over time among multiple items. It is this two-dimensional aspect of the comparison that makes the conventional bar charts inefficient in this case, and ternary plots are used instead. The latter provide an ideal way to depict compositional trivariate data and can be of use in many contexts. The structure of international migration flows by citizenship serves as an illustrative example only. I analyze immigration and emigration flows in the 28 European Union (EU) Member States (EU-28; as from 1.7.2013 till 31.1.2020) over the period 1998–2015. They are grouped into migration of nationals of a reporting country, (other) EU nationals, and non-EU nationals, which is a standard classification used in official migration statistics.

A web-based interactive visualization of ternary plots has been also developed to enable interested readers further exploration of the data (Nowok 2020). It complements the growing body of the online migration visualization tools that focus mostly on mapping bilateral (origin–destination) migration flows (see Dennett 2015 for a review), with only few examples that attempt to depict the complex pathways used by migrants (Allen 2018). It is also worth mentioning that a circular ideogram layout has been successfully adopted recently by migration researchers to present origin–destination flows, which can also capture the temporal data dimension in its interactive version (Sander, Abel, and Bauer 2014; Abel and Sander 2014).

## 2. Data on international migration flows

The underlying raw official data on international immigration and emigration is free and publicly available from the Eurostat’s online database (Eurostat 2020). It comes from two tables: “Immigration by age group, sex and citizenship (migr\_imm1ctz)” and “Emigration by age group, sex and citizenship (migr\_emi1ctz).” The EU and non-EU aggregates for the moving composition of the EU are published in the database, but they are incomplete. When an EU aggregate is missing but information on all required countries of citizenship is available, it is derived as a sum of its components, and a non-

EU aggregate is calculated as a difference between the overall migration total and an EU aggregate. However, this solves the missing data problem only partially, and many observations are unavailable.

There was no attempt to complete data using other sources. No data adjustments were made, and any interpretation should be made with great caution, because the quality of migration data, especially on emigration, is very often questionable. Besides, in theory, a (long-term) migrant is a person who moves to a country for at least a year, but migration definition may vary between countries and over time (Nowok, Kupiszewska, and Poulain 2006; Kupiszewska and Nowok 2008). Note also that due to later revisions some official statistics published by National Statistical Institutes may differ from the ones provided to the Eurostat and available in the Eurostat's online database.

### **3. Interactive ternary plots**

With an aim to illustrate and aid exploratory analysis of compositional data, an interactive web application has been developed (Nowok 2020) using R (R core team 2019) package shiny (Chang et al. 2019), where ternary plots are produced with ggtern package (Hamilton and Ferry 2018). Below a brief description of ternary plots is followed by an exemplary plot from the Shiny visualization and an overview of the app's features.

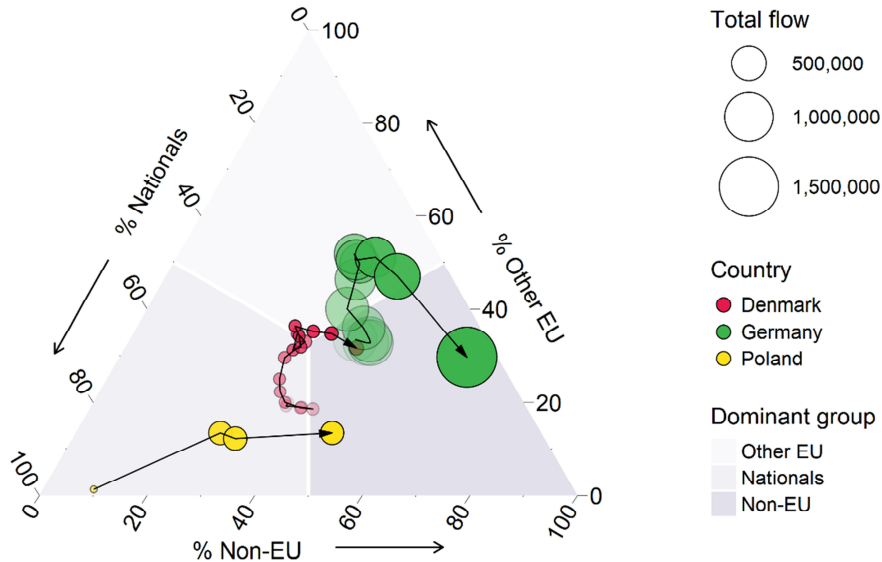
A ternary plot, also called a triangle or simplex plot, was specifically designed to depict compositional trivariate data (Aitchison 1986) and is commonly used in geology and other fields of physical science. It is similar to a scatterplot but depicts a closed three-component composition, with each point representing a unique composition of the three variables. The proportions of the three components plotted sum up to some constant that is represented as 1 or 100%. A ternary diagram has three axes, one for each variable, which form an equilateral triangle. The shortest distance from a point to the side opposite a vertex represents the value for the specific variable, so each vertex represents 100% of one of the variables, and the edge opposite represents 0%.

In the international migration example we have the following three categories: nationals of a reporting country, (other) EU nationals, and non-EU nationals. As an example, consider a composition of international immigration flow to Germany in the last observed year, 2015. It is depicted in Figure 1 by the darkest green circle (data points referring to earlier years are presented with higher transparency) and indicated also with an arrowhead (arrows connecting the circles aim to facilitate following developments over time). As regards the proportion of German citizens (nationals), the tick marks are on the left side of the triangle, and the lower left vertex represents 100%

share. The circle of interest is close to the opposite edge, equivalent to 0%, which means that the share of German citizens in 2015 was minimal (around 5%). The positioning towards the remaining edges tells us that the majority of migrants were non-EU nationals (around 65%), and other EU nationals constituted about 30% of the total flow. The background of the plot colored with different shades of grey helps to identify a dominant component of a migration flow. We can immediately see that for a few years immigration to Denmark was very balanced with equal share of each group (see the red points close to the triangle center).

As presented in Figure 1, data can be plotted for multiple states simultaneously. They are depicted with different colors, which allows to identify similarities and differences in migration patterns between them. Each circle represents migration flow in a single year. The area size of a circle stands for the total migration flow. Due to constraints in human perception, we are not good at matching sizes to counts, but we can still get an impression of magnitude and say which migration flow is bigger or smaller than another one (Meirelles 2013), which is sufficient here. Note that some of the features can be deactivated in the Shiny app, namely the path arrows and the size of the total flows (see Figure 2), which is advisable when data for many countries is displayed.

**Figure 1: International immigration to Denmark, Germany, and Poland by broad group of citizenship (nationals, other EU-28, and non-EU-28), 1998–2015**



Source: Eurostat, own calculations.

The shiny app includes data on both immigration and emigration. They can be displayed at the same time in a faceted graph, which facilitates comparison. Developments over time can be followed with an animation feature, but data can also be plotted for selected years only. Besides, the impact of the entry of new countries to the EU can be assessed by comparing data referring to the EU composition of the 28 Member States with data referring to the EU composition of the reference period. Data on total migration flows and proportions for the three citizenship groups that is used for plotting is displayed in a separate tab and can be consulted by a user if needed.

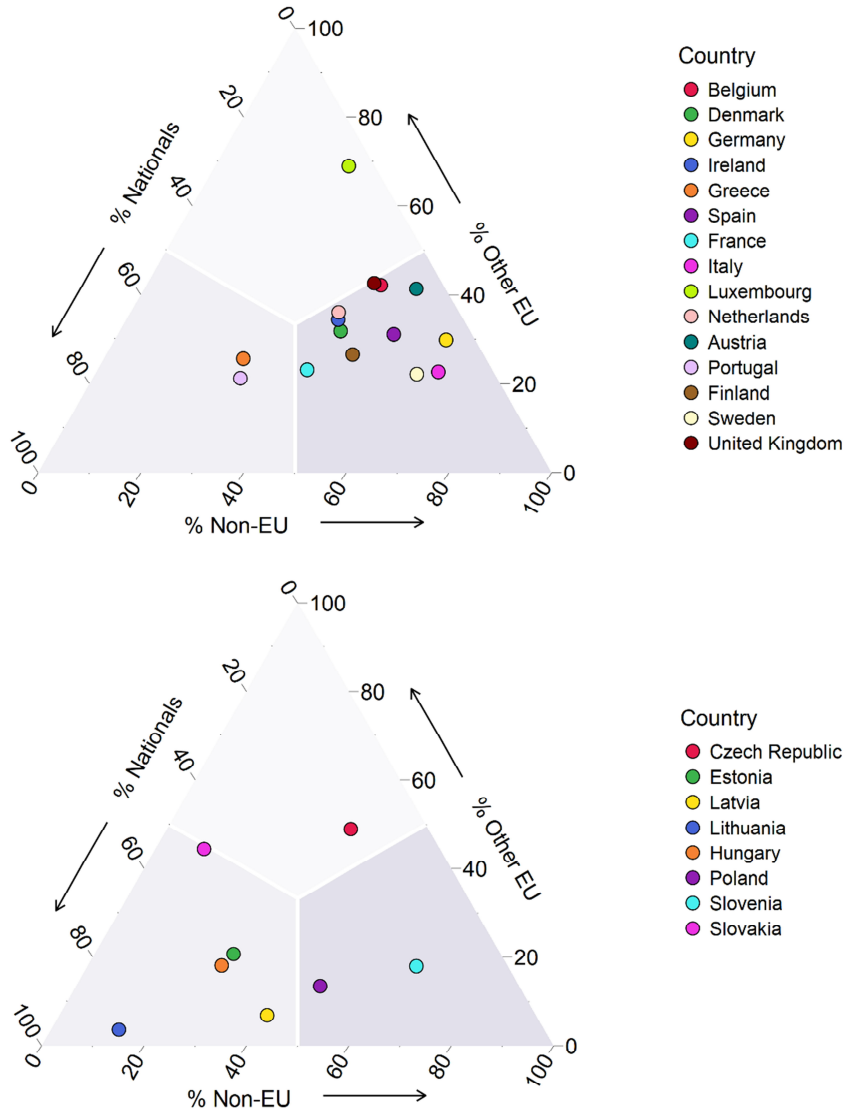
#### 4. Illustrative examples

One of the goals of using ternary plots is to compare many observations of compositional data in order to discover general patterns and spot unusual observations. Figures 2–4 presented below illustrate such comparisons. The presented diagrams can

be produced using the Shiny app, and to explore further details of the described examples, the reader is encouraged to navigate to the website (<https://bnowok.shinyapps.io/eumigration/>). Figure 2 shows international immigration in 2015 in the EU-15 (top panel) and in the EU-8 (bottom panel) by broad group of citizenship defined by the EU composition with the 28 Member States. The EU-15 represents the EU prior to the largest enlargement on 1 May 2004, and the EU-8 includes countries that joined the EU in 2004, excluding Cyprus and Malta. As we can see, the majority of the EU-15 states attract mostly migrants from outside the EU. In countries like Italy, Germany, and Sweden, their share exceeds 60%. Other EU-28 nationals constitute usually the second-largest group. For most countries their proportion varies between 22% and 43%. The most distinctive exceptions to the general pattern are Luxemburg, with a very large share of the other EU-28 nationals (69%), and Portugal and Greece, with a high proportion of nationals (around 50%). As expected, migration patterns in the EU-8 are quite different and more diverse. As opposed to immigration to the EU-15 countries, most flows are dominated by nationals of a destination country, and the share of other EU-28 nationals is lower than 20%. A relatively high share of immigrants coming to the Czech Republic (49%) and Slovakia (44%) from other EU-28 countries reflects migratory exchange between the two countries and an inflow mostly from outside the EU-15 (e.g., from Hungary, Romania, and Bulgaria).



**Figure 2: International immigration in the EU-15 (top) and the EU-8 (bottom) by broad group of citizenship (nationals, other EU-28, and non-EU-28), 2015**

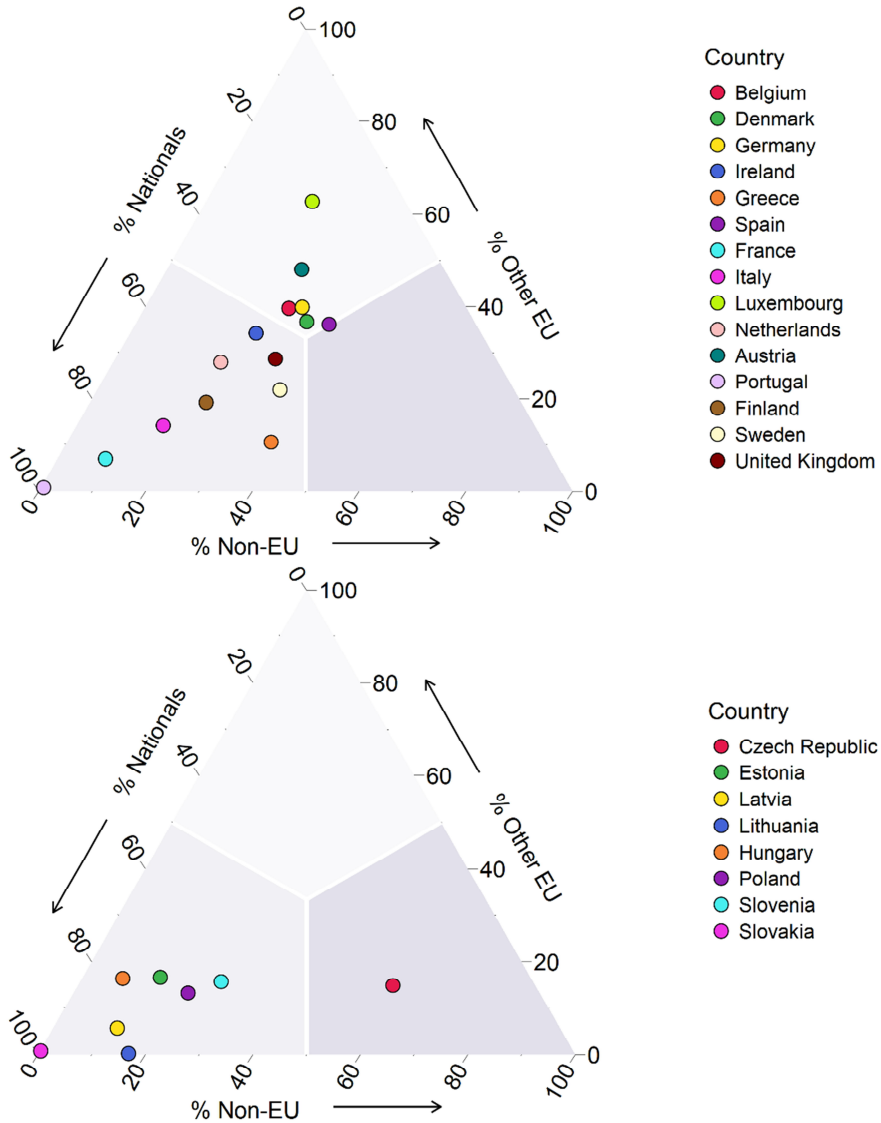


Source: Eurostat, own calculations.

Figure 3 presents similar data for emigration. We can see that the difference in composition between immigration and emigration is quite distinct. As regards emigration from the EU-15, there are a few countries that have quite balanced composition (the points close to the triangle center). There is also a clear general pattern of more or less equal share of other EU-28 nationals and non-EU-28 nationals (the former are dominating in some cases) with a varying share of nationals of the reporting country. Luxemburg, Greece, and Portugal are again exceptions. As regards emigration from the EU-8 countries, all outflows, except the one from the Czech Republic, are dominated by nationals of the reporting country. Nationals from other EU-28 states, whose proportion is nowhere larger than 17%, constitute the smallest group.

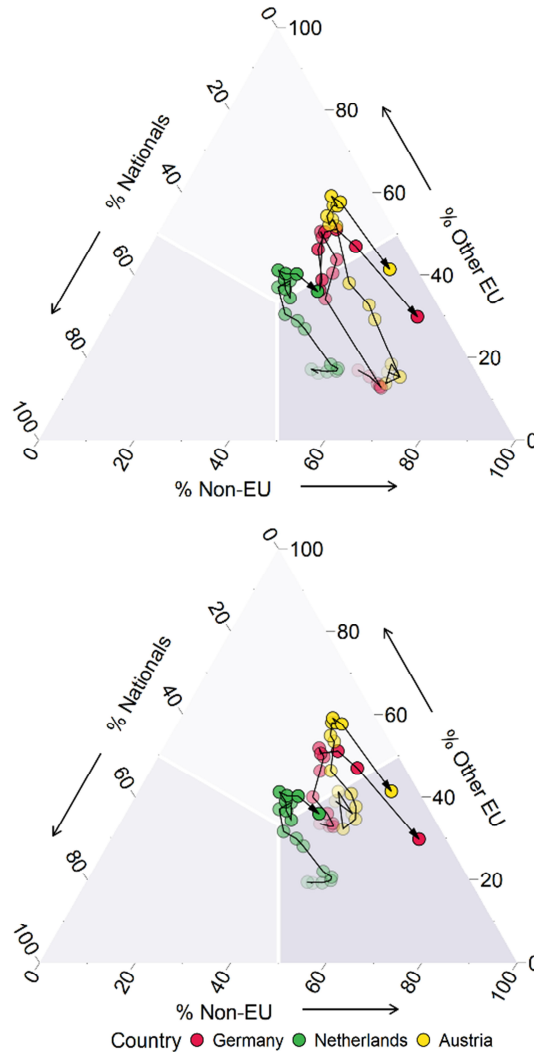
The last example illustrates similarities of changes occurring over time (1998–2015) in different countries (Germany, the Netherlands, and Austria) and the impact of the EU enlargement on the composition of immigration flows (see Figure 4). The comparison of the two panels shows that accessions of the new countries made a real difference for immigration to the Netherlands, but for Germany and Austria it mainly meant a reclassification of nationals of the new EU countries who were immigrating to these countries already before (data points at the beginning of the arrows moved along isolines for nationals, indicating change in proportions between non-EU and other EU nationals). In the following years the share of other EU nationals first continues to increase and then starts to drop, along with decreasing share of nationals. In 2015 there is a surge in the proportion of non-EU nationals, especially in Germany, which reflects the influx of asylum-seekers.

**Figure 3: International emigration in the EU-15 (top) and the EU-8 (bottom) by broad group of citizenship (nationals, other EU-28, and non-EU-28), 2015**



Source: Eurostat, own calculations.

**Figure 4: International immigration to Germany, the Netherlands, and Austria by broad group of citizenship for EU composition of the reference period (top) and EU composition with 28 Member States (bottom), 1998–2015**



Source: Eurostat, own calculations.

## **5. Concluding remarks**

An exploratory analysis of compositional data using ternary plots is restricted to a structure consisting of only three components, but possible applications are still numerous. We can focus on the most important triples of variables or group some variables and present aggregated percentages. Besides, additional dimensions of information can be represented using other visualization channels. Ternary plots have a number of advantages compared with bar charts. Trends and patterns can be found easily on ternary plots. Comparisons can be made at a glance, and there is practically no limit to the number of points that can be depicted in a ternary plot if we want to get a general insight into compositional variability. We can get an overview of the distribution of data, identify clusters, or add a series of contour lines to look at the density of a fourth variable of interest. Such plots can enhance exploratory analysis and communication of research findings, but because they are not commonly used in demographic research and public discourse, simple interpretation guidelines have to be set.

## References

- Abel, G.J. and Sander, N. (2014). Quantifying global international migration flows. *Science* 343(6178): 1520–1522. doi:10.1126/science.1248676.
- Aitchison, J. (1986). *The statistical analysis of compositional data*. London: Chapman and Hall. doi:10.1002/bimj.4710300705.
- Allen, W.L. (2018). *Representing freedom and force: How data visualisations convey the complex realities of migration*. Oxford: European Studies Centre (Dahrendorf Programme for the Study of Freedom Essays).
- Chang, W., Cheng, J., Allaire J.J., Xie, Y., and McPherson, J. (2019). shiny: Web application framework for R. R package version 1.4.0 [electronic resource]. <https://cran.r-project.org/package=shiny>.
- Dennett, A. (2015). Visualising migration: Online tools for taking us beyond the static map. *Migration Studies* 3(1): 143–152. doi:10.1093/migration/mnu073.
- Eurostat (2020). Eurostat's online database [electronic resource]. <http://ec.europa.eu/eurostat/data/database>.
- Hamilton, N.E. and Ferry, M. (2018). ggtern: Ternary diagrams using ggplot2. *Journal of Statistical Software, Code Snippets* 87(3): 1–17. doi:10.18637/jss.v087.c03.
- Kupiszewska, D. and Nowok, B. (2008). Comparability of statistics on international migration flows in the European Union. In: Raymer, J. and Willekens, F. (eds.). *International migration in Europe: Data, models and estimates*. Chichester: John Wiley and Sons: 41–71. doi:10.1002/9780470985557.ch3.
- Meirelles, I. (2013). *Design for information: An introduction to the histories, theories, and best practices behind effective information visualizations*. Beverley: Rockport Publishers.
- Nowok, B. (2020). International migration in the EU by broad group of citizenship (nationals, other EU, non-EU), 1998–2015 [electronic resource]. <https://bnowok.shinyapps.io/eumigration/>.
- Nowok, B., Kupiszewska, D., and Poulain, M. (2006). Statistics on international migration flows. In: Poulain, M., Perrin, N., and Singleton, A. (eds.). *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: Presses Universitaires de Louvain: 203–231.

- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [electronic resource]. <https://www.r-project.org>.
- Sander, N., Abel, G.J., and Bauer, R. (2014). The global flow of people [electronic resource]. <http://www.global-migration.info/>.
- Schöley, J. and Willekens, F. (2017). Visualizing compositional data on the Lexis surface. *Demographic Research* 36(21): 672–658. doi:10.4054/DemRes.2017.36.21.

*Nowok*: A visual tool to explore the composition of international migration flows in EU countries, 1998–2015