Forecasting Migration: The Role of Modelling Emigration – an Analysis for Germany

Johann Fuchs and Doris Soehnlein (both Institute for Employment Research, Germany)

Introduction

Many OECD countries are confronted with two population related trends. On the one hand, the labour force is expected to decline due to demographic ageing. On the other hand, there has been a tremendous inflow of migrants for a couple of years. Although migration raises some social problems, an example is the controversial public discussion in Germany following the mass inflow of refugees in the last few years, it could mitigate, in principle, an expected lack of labour and strengthen the sustainability of the social security system. Apart from any social or similar problems and from the question, where the migrants could come from in future, it seems to be a common consensus in the public as well as in the political discussion that at least a partial offset of the negative demographic trend might be possible by a certain level of immigration.

Actually, since the seminal study of the UN population division (2000), the term "replacement migration" has gained in some importance, even though it was neither the first nor the only study that deals with replacement migration. In addition, the UN report has been subject to some criticism (see Bijak/Kupiszewska/Kupiszewska 2008). Nevertheless, key results of the UN study are plausible and, from a political point of view, disillusioning. In order to keep the total population or some other demographic indicator constant, an extremely high level of migration would be needed to compensate for a long-lasting, negative natural population movement. Obviously, the scale of the necessary immigration depends on the indicator that should be kept constant (respectively almost constant). If one's interest, for example, is the size of the total population far less immigration will be required for a constant population stock compared to a policy aim to stabilize the old age dependency ratio (for an overview of the discussion, see Bijak/Kupiszewska/Kupiszewska 2008). In the light of high immigration, these reservations against replacement migration seem to be less important.

In the political realm, moreover, these expectations are often based on current migration figures. In addition, it looks as if official population predictions support these expectations because their underlying assumptions are also influenced by recent events. The Federal Statistical Office of Germany (Destatis), for example, produces population forecasts using migration scenarios which are usually closely oriented toward the most recent net migration flows. Although Destatis offers a wide variety of scenarios, which are quite different with respect to the underlying net migration, there is always a favored scenario that is based on recent developments. Insofar, the typical approach of an official forecast can undergird and even reinforces the expectation of a high yearly net migration over the coming decades.

We see two shortcomings of this view, which are related to each other: firstly, the public discussion does not differentiate migration balance and gross migration flows; secondly, the role of emigration is ignored. In fact, the number of emigrants could be expected to rise with the growth of the domestic population. A higher level of immigration increases the population and, therefore, more people are under risk to leave the country. Accordingly, either the emigration rate drops or the annually immigration inflow has to increase in order to achieve the necessary net inflow.

Our study focuses on this link and answers the question to what extend the levels of migration flows, the net migration and the emigration rate, respectively, are effected by the way migration is modelled. The analysis should allow an evaluation of current population respectively migration projections as well as a better understanding of the possibility of replacement migration. Hence, we add results for some selected population indicators, e.g. the working age population.

To determine concrete figures for our question, migration has to be embedded within a population model, i.e., the other demographic components, fertility and mortality, has also to be considered. However, our paper does not discuss the question of “forecast”, “projection” or whatever term. We understand the following estimations more or less as a kind of simulation. Various simulations show how the results of a typical population projection change with the underlying migration model.

Current practices of forecasting migration - a brief review

The methodology to forecast the future population (including its structure) can differ in various aspects. Alongside with the basic question, whether the model should follow a deterministic or a stochastic approach (for a comprehensive discussion of this issue, see e.g. Lee/Tuljapurkar 1994, National Research Council 2000, p. 190f.), previous studies differ particularly with regard to the migration modelling. Bijak (2012) compiled an overview of the current practices of the statistical offices for selected countries in respect to the migration assumptions. He listed inter alia Australia, France, Germany, Italy, Japan, Spain, the UK, the USA and Eurostat.[1]

---

[1] Sometimes the national publications are a bit vague, so it is not quite clear in all cases what exactly was projected. Bijak, for example, classified Germany to the net migration countries, as almost every publication of the Federal Statistical Office (Destatis) refers to net migration flow only. In fact, Destatis (2015) set out separate assumptions for immigration and emigration. According to personal communication with Destatis, they assume 600,000 emigrants per annum for base migration.

Looking at the list of Bijak (2012), augmented by own enquiries in respect to recent non-official studies, we see many forecast approaches focus on net migration. Modelling net migration has the great advantage of being simple and even possible in case of limited data. However, it has some flaws. In case of negative net migration figures for some age-sex combination, negative cells of the population stock are possible. In addition, this model does not tell us anything about the numbers of immigrants and emigrants, including their different age profiles, which might be useful information.

Modelling separately gross flows of both immigration and emigration overcomes some drawbacks of the net migration type. In this case, the projection specifies a certain number of immigrants as well as emigrants. (It should be mentioned that, typically, several variants with alternative assumptions in respect to the levels are calculated.) The emigration level works as a “base” emigration if a certain volume of net migration is intended. This is the approach Destatis currently applies for Germany. Modelling migration separately on a level basis provides information about volumes and structure, and allows for a more sophisticated modelling of the age profiles, e.g. the use of schedules (Raymer/Roger 2008). Nevertheless, negative cells of the population stock are still possible.

A less frequently applied approach is a mixture of a gross immigration flow and an emigration rate. Bijak (2012), who preferred rates in general but admitted that immigration rates are hard to obtain, recommended this procedure as a compromise. This approach at least ensures that emigration will never exceed the population in any age group, in other words, it prevents the possibility of a negative population stock.

Current studies for Germany, which is a country with strong immigration flows in the recent past, a long-lasting demographic ageing and several signs of skill gaps now, can serve as an example, how these different approaches work and how different the results are with regard to the predicted net migration. Most demographers expect Germany to have a high migration inflow in the future.

Applying a stochastic population forecast, Deschermeier (2016) modelled net migration from 2016 to 2035. He yielded an average net migration of about 218,000 persons annually. Destatis updated in 2017 its previous long-term population forecast, the 13th coordinated population projection, which had been published in 2015. This was done in the light of the refugee crisis and high immigration from EU-countries over the course of the financial and economic crisis in several European countries. The updated projection assumes a net migration of 300,000 p.a. up to 2060 (Destatis 2017). Recently, Destatis (2019) published its 14th projection, covering 27 variants plus three variants with special assumptions. This release includes three migration assumptions, which range from almost 150,000 net migration per year to just above 300,000 migrants. As already mentioned, Destatis separates immigration and emigration, assuming a base migration of 600,000 emigrants. The stochastic model of Fuchs et al. (2018) distinguishes between German and non-German nationals and estimates the inflow of immigrants on level-basis and emigration by group specific emigration rates. This study produced an annual median net migration of less than 150,000 non-German migrants on average over the forecast period 2016 to 2060. Vanella/Deschermeier (2018) stochastically estimated age-, sex- and nationality-specific net migration up to 2045. Median net migration was predicted to be more than 250,000 migrants over the complete forecast horizon.

The cited studies differ in various aspects, for example with regard to the initial year. As far as we can conclude from the details, however, the most important difference is the way migration is modelled. In particular, we know from the population statistic for Germany that the overall emigration rate proves to be rather stable. In case of the non-nationals, it remains on a long-run level of about 8 percent of the population. We assume a feedback mechanism caused by the emigration rates that influences the results, as we observe a positive correlation between emigration and past immigration flows. Lagged one year, emigration correlates with immigration by 0.9 over the last 60 years. Therefore, this brief overview suggests having a closer look at the concrete modelling approach.

Data and method

Our analysis is based on the most up-to-date population figures, predominantly available on Genesis Online provided by Destatis. The data are structured by single year of age, sex, and nationality (German nationals, non-German nationals).

We apply a typical cohort-component approach to project the population development. As we are interested in the effect of a different modelling of migration, we specify a deterministic simulation model. Assumptions in respect to fertility and mortality follow the projection of Destatis (2015). Fertility and mortality rates are held constant for all simulations.

The migration part of the model differentiates between immigration and emigration, both distinguishing between German and non-German nationals.

Immigration flows were estimated on a level basis. For the purpose of comparison, we calculated several scenarios with different immigration levels, keeping the age and sex structure constant.

For the emigration flows, two simulation models are applied. One is based on a gross flow assumption, applying a constant distribution of the gross flows among age and sex. For this study, we assume a yearly emigration of 600,000 non-German citizens. The second approach takes emigration rates to estimate the emigration numbers. Hence, total emigration ceteris paribus increases if the population grows by immigration.

The emigration rates have to be projected for all ages by both sexes and both nationalities. A naive forecast can use constant rates. We rely on this approach for our base simulation. In addition, we apply a more sophisticated model, conducting a principal components analysis (PCA). PCA reduces the high dimensionality and accounts for autocorrelation and the cross-correlation between the rates. For our study, we calculate the principal components (PCs) for both sexes independently. PCs with an eigenvalue above one were forecast using time-series methods. The future age and sex specific emigration rates finally relies on the predicted PCs and stochastic as well as deterministic trends.

The simulation horizon is 2060, which allows comparing our results with other long-term projections. This time horizon also covers conceivable longer-term developments, e.g. long-run baby boom effects, which could influence the emigration numbers.

To simplify matters, only the emigration rates of the non-German population are modelled. Migration flows of the population with German nationality are comparatively small (net migration 2010-2015 was 0.04% of the population) and in the past net migration was often close to zero (-18,000 in 2015, compared to a population of 55 million German nationals). Therefore, we decided to set out both emigration and immigration of German nationals to zero over the simulation period for this simulation study.

First results

Our paper focuses on the resulting migration variables, i.e. net migration, gross immigration and emigration and the total emigration rate. Further demographic indicators, like the total and the working age population, the old age and the total dependency ratio will complete the picture.

First estimates show the expected result that a comparatively huge immigration is necessary to keep the change of the selected demographic indicators within a visible range. Furthermore, immigration has to grow to – over the long run – unprecedented numbers in case the emigration rate should not break off totally. From the opposite perspective, an almost stable emigration rate in connection with visible immigration numbers restricts the net migration surplus.

How the use of predicted emigration rates change the results, is still work in progress. The fixed rates model, with immigration rates from the year 2012, and the stochastic rates model essentially differ at the beginning of the simulation period.

Preliminary conclusions

In the light of the first results, further methodologically extensions are envisaged. Although the rates models incorporate the influence of age, sex and (partly) nationality, they ignore other important influential factors that governs emigration. Well-known factors, like the labour market performance (employment/unemployment, wages) or the average duration of staying in the country (e.g. Hansen 2013), would be nice to become implemented in the rates model.

Some of the factors, however, can hardly be predicted in the long run. Forecasting labour market variables is at least a similar challenging task as forecasting migration. Others, like the duration, are currently not available for Germany in the necessary manner. Moreover, we anticipate an endogeneity problem.

Depending on the data availability and quality, it might be possible at some point in the future to improve the analysis by a breakdown of the data due to the migration intentions, e.g., for circular migration, flight migration and other reasons. These intentions has to be embedded in a model that forecast immigration in the same structure and could proxy some results from life cycle models (Borjas/Bratsbert 1996, Burkhauser et al. 2016).

Improving the short-run performance, e.g. by applying a functional data approach, is extremely necessary and should at least be possible. We are still working on it.

Our results show that in Germany an almost stable share of the non-German population leave the country each year. More immigration, therefore, would ceteris paribus cause more emigration. An important preliminary political conclusion is that Germany, and presumably many other OECD countries as well, should do better in reducing emigration by promoting those migrants who already live in the country, rather than attracting additional migration inflow.