A Probabilistic Cohort-Component Model for Population Forecasting – The Case of Germany

Patrizio Vanella
Gottfried Wilhelm Leibniz Universität
Center for Risk and Insurance
Otto-Brenner-Str. 1, DE-30159 Hannover
pv@ivbl.de

Philipp Deschermeier
Institut Wohnen und Umwelt
Field of Housing
Rheinstr. 65, DE-64295 Darmstadt
p.deschermeier@iwu.de

Abstract

The future development of population size and structure is of importance since planning in many areas of politics and business is conducted based on expectations about the future makeup of the population. Countries with both decreasing mortality and low fertility rates, which is the case for most countries in Europe, urgently need adequate population forecasts to identify future problems regarding social security systems as one determinant of overall macroeconomic development. This contribution proposes a stochastic cohort-component model that uses simulation techniques based on stochastic models for fertility, migration and mortality to forecast the population by age and sex. We specifically focus on quantifying the uncertainty of future development as previous studies have tended to underestimate future risk.

The results provide detailed insight into the future population structure, disaggregated into both sexes and 116 age groups. Moreover, the uncertainty in the forecast is quantified as prediction intervals for each subgroup. The underlying models for forecasting the demographic components have been developed in earlier studies and rely on principal component time series models. Since the proposed model is fully probabilistic, it offers a wide range of information, not only identifying the most probable course of the population but also a vast number of possible scenarios for future development of the population and quantifying their respective likelihoods. The model is applied to forecast the population of Germany until 2040.

The results indicate a larger future population for Germany compared to the population predicted in studies conducted before 2015. The driving factors are lower mortality, higher fertility and higher net migration as derived by us statistically in contrast to widely used qualitative assumptions. The present study shows that the increase in population is mainly due to a larger proportion of older individuals.

Keywords: Population Forecasting; Stochastic Simulation; Cohort-Component Methods; Principal Component Analysis; Time Series Analysis

1 Introduction

The future development of the population structure is of immense importance since planning in many areas of politics and business is done based on expectations about the future composition of the population. Countries with low fertility and decreasing mortality rates, as is the case for most countries in Europe, particularly need accurate population forecasts since these demographic changes transform the long-term age distribution of the population in favor of older persons. These changes result in widely discussed future problems, e.g., for the social security systems as well as the labor market as a whole. The public discussion about the demographic change in Germany and its challenges is mostly tinged with negative undertones (Deschermeier 2011, 669). Nevertheless, the transformation of a society also represents a very positive aspect: people are getting older while experiencing more healthy and active years of life compared to those in previous generations (Schnabel et al. 2005, 3).

The shrinking and aging of the population in Germany in the long run was considered consensus by experts (Fuchs and Dörfler 2005, 1, Lipps and Betz 2005, 32, Eisenmenger et al. 2006, 34, Börsch-Supan and Wilke 2009, 32, Härdle and Myšičková 2009, 26, Destatis 2009, 12, Dudel 2014, 184, Deschermeier 2015, 106, Pötzsch and Rößger 2015, 15, Fuchs et al. 2018, 48-9). Deaths have exceeded births in Germany since 1972, so without positive net migration the population would be shrinking (Swiaczny 2016, 158). Between 2009 and 2015, net migration into Germany has increased monotonically, starting at a level of almost -56 thousand in 2008 and reaching a record of more than +1.139 million in 2015 (Bundesministerium des Innern 2017, 186). This trend was first induced by a combination of three factors. First, the European debt crisis, which hit countries in Southern and Eastern Europe especially hard and led to major immigration from these regions to Germany and to Central and Northern Europe (Brücker et al. 2017a, 3). Second, a large increase in immigration from Afghanistan caused by a worsening security situation due to increased aggressiveness by the Taliban against the American military and civilians (Bundesministerium des Innern 2011, 107, International Organization for Migration 2014, 104). Third, a large spike in migration from Iraq following the start of the resurgence from the United States *(U.S.)* military, which caused more attacks from Islamist militias (Bundesministerium des Innern 2011, 107, Jaffe 2009). These trends continued in the following years, and the enlargements of the European Union *(EU)* led to an increased influx of people from Southeastern Europe to Germany (Bundesministerium des Innern 2015, 14). Following the so-called *Arab Spring*, which started in 2011 in Tunisia, Islamists have gained massive amounts of power due to the power vacuum appearing after the end of dictatorships in these countries (Council on Foreign Relations 2017). The so-called *Islamic State (IS)* in 2014 had rapid and surprising military success, especially in Syria and Iraq, where they proclaimed a caliphate (Heidelberg Institute for International Conflict Research 2017, 189). Many people subsequently fled from these regions, leading to record refugee migration into Germany in 2015. This migration was fueled by Chancellor Merkel's decision to simplify the asylum process and to essentially guarantee people from Syria legal refugee status. These changes subsequently motivated many people in places such as Serbia, Albania, Kosovo and Iran to try their chances as refugees as well. Some of these refugees even immigrated illegally using false identification documents to pose as Syrians (Aust et al. 2015, Bewarder and Leubecher 2016, Bundesamt für Migration und Flüchtlinge 2016, 14-50, Zeit Online 2015).

These developments will probably have a long-term effect on the level of migration to Germany as well as the fertility level because migration from countries in the Middle East may influence the reproductive level in Germany (Milewski 2010, 310-6, Naderi 2015, 329). Therefore, the demographic outlook in Germany has changed substantially due to the events occurring during the past decade.

Against this background, this contribution provides a stochastic population forecast of the year-end population in Germany through the year 2040. The population in each year of the forecast is broken down by sex and age for the range 0 to 115 years. We use stochastic modeling approaches developed in past contributions (Vanella 2017a, Vanella and Deschermeier

2018a, 2018b) to forecast the demographic components of the population development. We use these forecasts to estimate the growth in the age- and sex-specific population, starting from the estimated population on December 31, 2016. In this way, we generate sample paths for the future population by simulating a probabilistic cohort-component model. The underlying concepts may also be used to forecast populations in other countries or other regional units.

Stochastic approaches are gaining popularity as an alternative to the common deterministic population projections that use scenarios to address future uncertainty (Keilman et al. 2002, 410). The ability to provide a measure on the entry probability is key for the users of these forecasts. Stochastic forecasts based on simulations are less prone to subjective decision-making. Our model returns not only the median age- and sex-specific population up to the year 2040 but also quantifies the uncertainty in the forecast, with 75% and 90% prediction intervals (*PIs*) for each year, age and sex.

The next section presents a condensed historical overview of the evolution of the cohort-component method for population updating and past advances in population projection, starting with the first deterministic models and continuing with improvements to these models through probabilistic forecasting. Our study primarily focuses on Germany; therefore, our overview is essentially limited to population projection in Germany. In Section 3, we describe the population forecast process in detail by explaining how the demographic components fertility, migration and mortality are forecast and how these individual forecasts are combined into an overall population forecast for Germany via a probabilistic cohort-component model. Section 4 presents and discusses the results, and Section 5 provides an outlook and discusses the limitations of the presented approach.

2 Selected Population Forecasts and Projections with special Emphasis on Germany

Future population projections for Germany are often conducted by deterministic cohort-component models. To the best of the authors' knowledge, this method dates to 1863, when the Census Bureau of England and Wales (1863, 5) ran a projection of the population in England

and Wales for the year 1881 by 20-year age groups. Births, deaths and migrations were identified as the components of demographic development. The population was projected by making assumptions about the changes in birth rates, mortality rates and net migration for each age group, or cohort. Cannan (1895, 508-15) further developed the method by taking ten-year age groups and assuming trends in age-specific fertility derived from recent census data. He projected the population in England and Wales until the year 1951. Whereas Cannan's approach implicitly modeled international migration in combination with deaths, Whelpton (1928, 255-70) incorporated expectations of migration in a forecast of the U.S. population by age group, sex and ethnicity until the year 1975, setting the stage for modern cohort-component modeling.

Projections for the future population in Germany are often conducted using deterministic cohort-component methods. Deterministic methods quantify a limited number of scenarios whose likelihoods of occurrence are not quantified by probability. Therefore, stochastic methods are recommended for population forecasting (Alho and Spencer 2005, 2-3, Bomsdorf et al. 2008, 125, Keilman and Pham 2000, 42-3, Keilman et al. 2002, 410-2, Lee 1998, 157-70, Lee and Tuljapurkar 1994, 1175, Lutz and Scherbov 1998, 83). Ledermann and Breas (1959, 637-81) proposed the transformation of age-specific mortality rates *(ASMRs)* into indices through singular value decomposition, which was developed geometrically by Pearson at the beginning of the 20th century (1901, 559-63). They were thus the first to use principal component analysis (*PCA*) to reduce the high dimensionality in demographic processes. Le Bras and Tapinos (1979, 1405-49) elaborated on the preliminary work of Ledermann and Breas 20 years later in the first principal component (*PC*)-based population projection for France until the year 2075.

Bozik and Bell (1987) proposed a groundwork for stochastic modeling by applying autoregressive integrated moving average (*ARIMA*) models to forecast age-specific fertility rates *(ASFRs)* in the United States. Bell and Monsell (1991: 156-7) applied this method to forecasting age-specific mortality rates *(ASMRs)*. Lee and Carter simplified the Bozik-Bell and Bell-Monsell approaches to forecast age-specific mortality (Lee and Carter 1992, 660-8) and fertility rates (Lee 1993, 190-9) in the U.S. Since then, various modifications of the *Lee-Carter model* have been proposed (see, e.g., Booth 2006, 554-62, Booth et al. 2006, 290-304 for an extensive overview), maybe most notably the functional principal components approach of Hyndman and Ullah (2007, 4945-52).

Many population projections and forecasts[1] have been made for Germany during the past half-century; the best known is the "*koordinierte Bevölkerungsvorausberechnung*" from the German Federal Statistical Office (*Destatis*). The first version was published in 1966. Since then, twelve updates have been made with improved techniques. The basic principle involves making a set of assumptions about the long-term development of life expectancy, total fertility rate (*TFR*) and net migration (currently two alternatives for each) to derive age-specific statistics. These different assumptions are combined to create eight different realistic scenarios for the course of the future population until the year 2060 (Pötzsch 2016, 37, Pötzsch and Rößger 2015, 7-41). Similar methods are used by the European Union (2015, 14-29).

Probabilistic population forecasts for Germany are rare. To the best of the authors' knowledge, the first approach was undertaken by Lutz and Scherbov (1998, 83-91), although their effort may be considered to be a stochastic projection rather than a forecast. Their idea was to pool a large number of earlier deterministic projections and to approximate the distributions of the parameters by assuming Gaussian distributions. Lutz and Scherbov investigated nine population projections for Germany and derived distributions for the TFR, life expectancy and net migration. On the basis of these summary statistics and assumptions about the distributions of the age-specific rates, they calculated empirical quantiles for the population size via scenario-based simulation to obtain projection intervals through 2050. Although this method may be appropriate when lacking a sufficient statistical basis for inference, it is quite subjective since it is built upon the scientists' assessment of the future course of the demographic components. Subjective judgment generally has a high potential for error since it is not necessarily connected to statistical data. Furthermore, individuals experience difficulties in translating their qualitative judgment about realistic future scenarios into quantitative probabilities (Lee 1998, 168-70).

Lipps and Betz (2005, 11-38) produced separate forecasts for the population in West and East Germany for the period 2002-2050, assuming convergence of the mortality and fertility rates

---

[1] For further reading on the distinction between forecasts and projections, see e.g. Bohk 2012, 21-5.

in the East towards the levels in the West. They simulated 500 trajectories for a mortality index, the TFR and net migration. The age-specific mortality rates were derived through the classic Lee-Carter index, and the TFR was assumed to follow a *random walk process*[2]. Age-specific fertility rates (*ASFRs*) were deduced from the TFR with a variable Gaussian ASFR distribution. The net migration was modeled as an autoregressive process of order one (*AR(1)*). Age-specific migration was then calculated via a distributional assumption. The simulation of the time series processes produced 500 trajectories with PIs of the age- and sex-specific populations of West and East Germany.

The contribution followed some good approaches but was not flawless. The assumption of an age schedule for ASFR ignores the timing effect in reproductive behavior. Furthermore, the assumption of convergence of the mother's mean age at birth is restrictive and, at least from today's perspective, is not realistic at a level of 31.45 years[3]. Quantification of the PI for this statistic seems problematic since the variance in the forecast is apparently constant and has the same value for 2002 and 2050, which is implausible since uncertainty about the far future is logically greater than that for the near future. The random walk hypothesis for the TFR in general does not perform well in forecasting, as shown by Vanella (2017b, 25-6).

Bomsdorf, Babel and Schmidt (2008, 125-8) used ARIMA models to forecast the TFR and the net migration in Germany. They used these summary measures to derive ASFRs and age-specific migration via age schedules, namely, a Beta distribution for the ASFRs. Age- and sex-specific measures for mortality and net migration were obtained from the Lee-Carter model, and 5,000 simulations of the time series models produced empirical PIs. Härdle and Myšičková (2009, 4-26) applied the Lee-Carter models for mortality and fertility to estimate these two components for Germany. Furthermore, they forecast immigration to and emigration from Germany with separate AR(1) models to estimate the population in Germany until the year 2057.

Dudel (2014, 95-216) non-parametrically forecast the population of West and East Germany until 2060 using historical simulation techniques based on 1,000 trajectories. His method, although statistically interesting, has a few caveats. First, the mortality model assumes a perfect

---

[2] See, e.g., Dickey and Fuller 1979, 427, Vanella 2018, 230 for a definition of a random walk.

[3] The mean age at child birth in 2015 was 31, and long-term increases were nearly linear per annum for almost two decades (see GENESIS-Online Datenbank 2018c).

correlation between the two genders, which statistically is unlikely (see e.g., Vanella 2017a, 543-52). Although different developments in mortality are evident for both sexes, mostly arising from different smoking (Pampel 2005, 461-3, Trovato and Lalu 1996, 31-5, Waldron 1993, 458-60) and nutritional (Luy and Di Giulio 2006, 1-8) behaviors, the main trends in mortality reduction result from advances in medicine and better education among the population with regard to health and hygiene. Females and males both benefit from these improvements (Pötzsch and Rößger 2015, 34). Second, Dudel rejects trajectories for the TFR under 1 and over 3. The need for such strong restrictions implies that the method is not completely suited for the problem. Third, the overall migration model can be criticized because it assumes a fixed age schedule (which is unlikely) and PIs whose width remains almost constant over time instead of increasing, as should be the case for the far future. One may argue that a forecast horizon of 2060 is too long for estimating valid PIs, considering the relatively short base data, which leads to systematic underestimation of the future uncertainty of the prediction. Moreover, 1,000 trajectories is a low number for reaching convergence in the simulations (Vanella 2017a, 546).

Deschermeier (2015, 2016) forecasts the total population of Germany until 2035. He used a model designed by Hyndman and Ullah (2007) to forecast the ASFRs and applied an advanced version of Hyndman et al. (2013) to forecast ASMRs and net migration. Although the model appears plausible, it also underestimates the uncertainty in the forecast. Hyndman's approach smooths the data against outliers, which may be reasonable in some cases to obtain better estimates for the mean prediction. The problem with this method is that this smoothing ignores the probability of future outliers and therefore effectively underestimates the future uncertainty by simply stating that already observed outliers cannot appear again in the future.

The United Nations (*UN*) apply the Bayesian hierarchical model of Raftery et al. for quinquennial life expectancy and TFR projections for all countries (Alkema et al. 2011, 818-29, Raftery et al. 2013, 780-6, 2014a, 60-5, 2014b, 801-6, United Nations 2015, 15-33, 2017). Migration is not addressed stochastically, which results in significant underestimation of the uncertainty in the population projections in this case as migration is the biggest source of uncertainty. The UN emphasizes that this is a projection, not a forecast, and that the result should be interpreted cautiously. With a time horizon of 85 years and the available data history, forecasting the future with adequate risk does not seem reasonable. Nevertheless, projections with such a long horizon may be useful as well for estimation of possible future development under specified circumstances.

We derive the future uncertainty in a forecast using past data. With a base period of 20 years and a forecast horizon of 50 years, e.g., we would implicitly assume that only the trends observed during the last 20 years can happen over the next 50 years. Even with expert knowledge, this leads to the problem of objective and quantitative transformation of qualitative assessment into probabilities, as stated above.

Fuchs et al. (2018, 44-54) forecast the population until the year 2060 using time series methods for the principal components of the demographic rates. This method is the most convincing population forecast in Germany to date, but they also appear to underestimate the uncertainty because the PIs of the TFR and net migration remain essentially constant after 2020. Furthermore, the course of net migration is implausible because it decreases considerably until approximately 2020 and then increases slightly. Considering the high stochasticity of international migration (see Vanella and Deschermeier 2018a), proposing some assumptions on the future course is understandable, especially regarding the long-term convergence of net migration. A decrease in net migration by approximately 750,000 from 2015 to 2016, as assumed in the mean, has never been observed for Germany to our knowledge. Furthermore, the assumption that net migration will increase again after such a heavy decrease appears unrealistic.

A general problem of the mentioned studies is underestimation of the future risk in the population forecasts, as shown. Some models quantify uncertainty qualitatively, which is very difficult to translate, as shown earlier. On the other hand, the presented quantitative studies mostly use the Lee-Carter model for forecasting, which is mostly sufficient for the mean but naturally leads to underestimation of future risk due to omission of random variables from the analysis. The UN model does not quantify the uncertainty in migration at all.

3 Method and Data

In this section, we propose a population forecast based on a probabilistic cohort-component model. The partial models for the demographic components shall be explained shortly.[4] First, the age-, sex-, and nationality-specific net migration figures are forecast as in Vanella and Deschermeier (2018a). The data used are synthetic net migration figures per years of age (0-105), sex (binary) and nationality group. The nationalities are split into seven groups: Germans, EU- or Schengen-citizens excluding Germany, Third-Country Europeans, Africans, Asians, Citizens from the Americas or Oceania *("Overseas")*, and finally persons with no clear information on their citizenship, either because it is unknown or they have none *("NA")*.

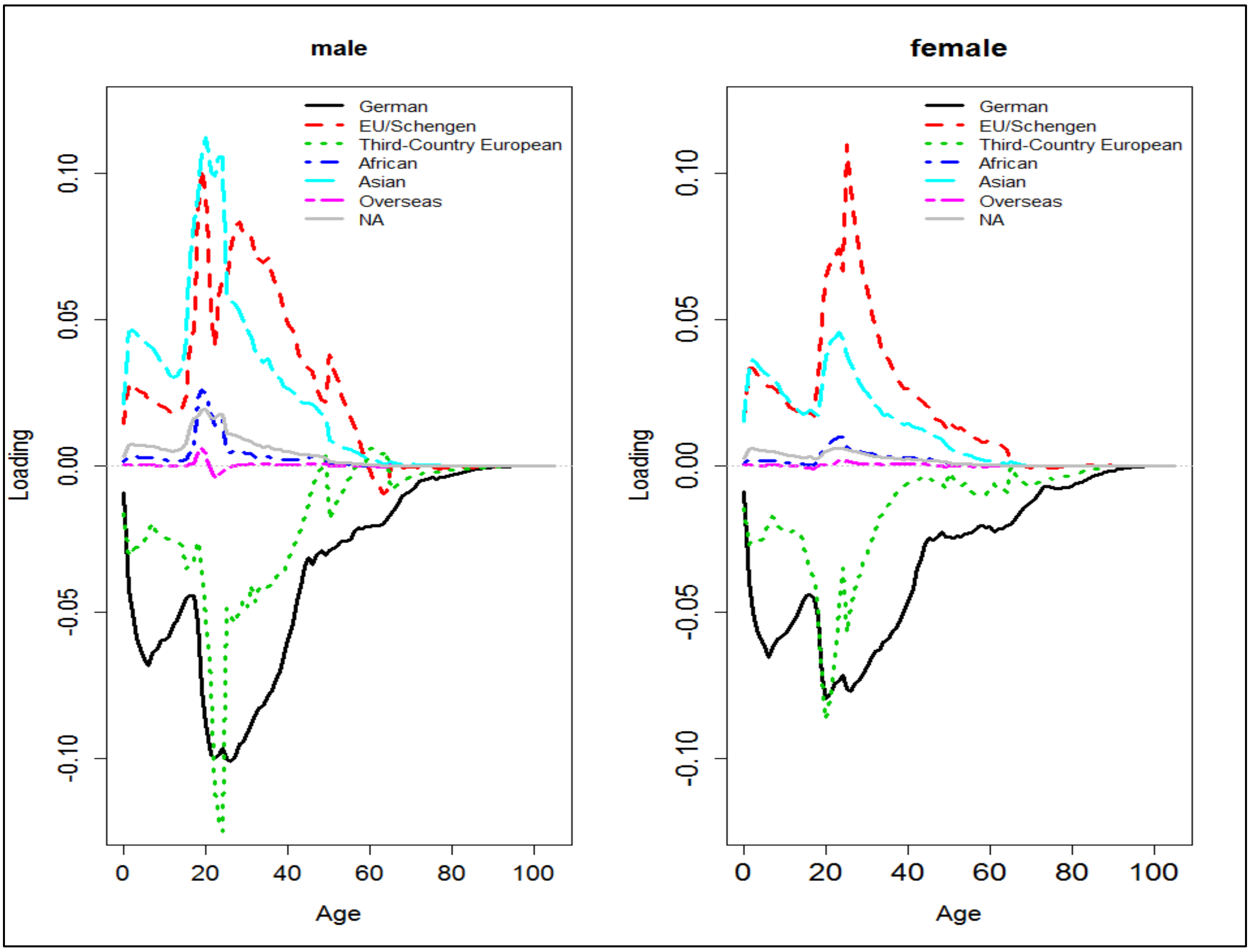
The dataset was estimated through two datasets provided by Destatis; the first includes age-specific migration data by sex, divided by Germans and non-Germans (Destatis 2015a, 2016a, 2017a, 2018a), and the second dataset is disaggregated by exact nationality and five age groups (Destatis 2017b, 2018b).[5] The base time period is 1990-2016. We run a principal component analysis (*PCA)* on the derived 1,484 age-, sex-, and nationality-specific net migration *(ASNSNM)* figures. The first two principal components *(PCs)* were identified as some kind of labor market index and an index for crises. The loadings of the PCs are for both sexes and the different nationality groups given in Figure 1 and Figure 2.

---

[4] The original sources serve as a more detailed description of the models and their results.

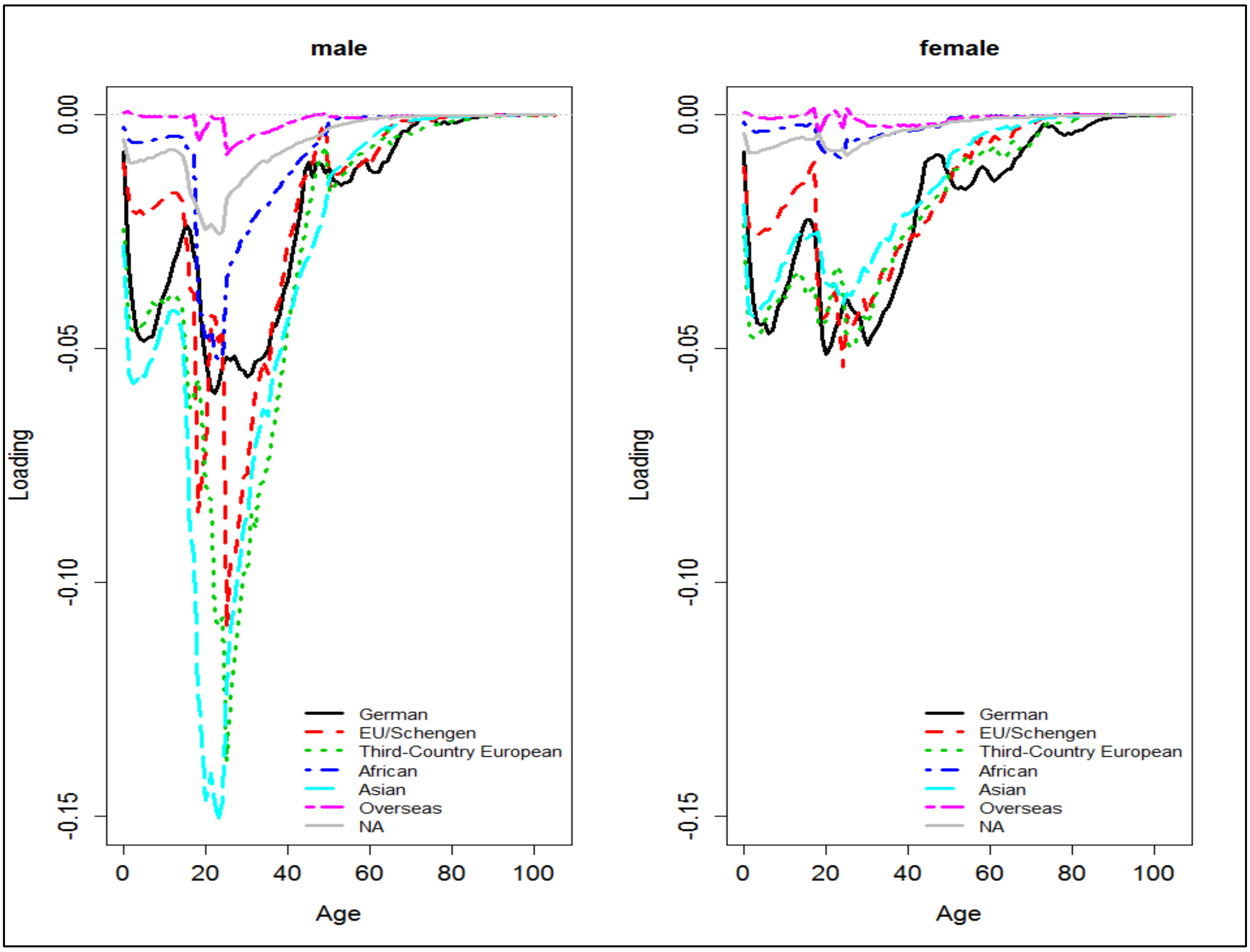
[5] The exact method for deriving the synthetic data is outlined in Vanella and Deschermeier (2018a, 264-71).

Figure 1. Loadings of the Labor Market Index



Source: Own calculation and design

Figure 2. Loadings of the Crisis Index

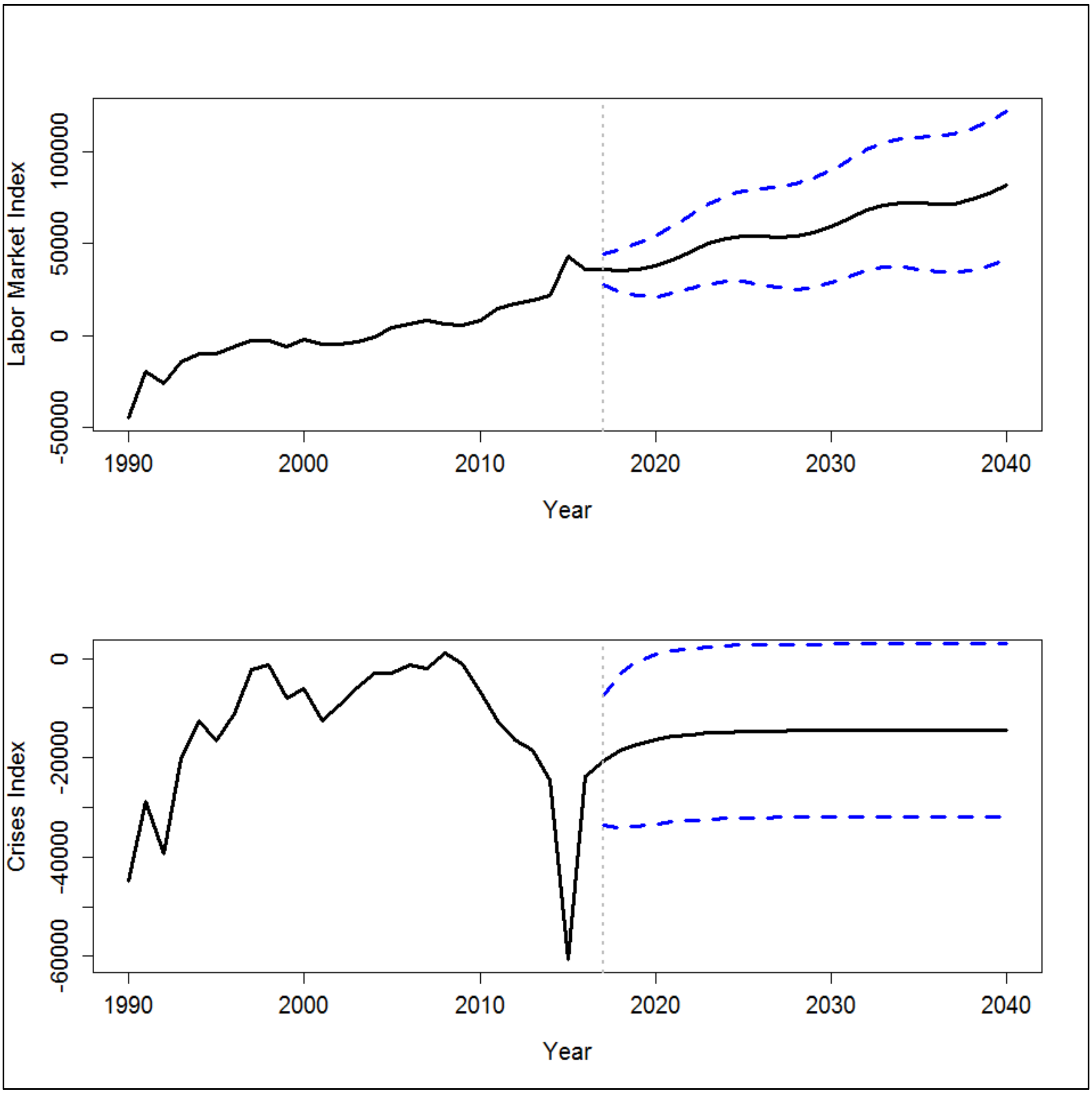


Source: Own calculation and design

The historical course together with the forecast of these two variables through 2040 is plotted in Figure 3. The Labor Market Index has an increasing long-term trend on average and includes cyclical effects, which are typical for labor markets. The Crises Index is assumed to converge towards its mean during the base period due to a lack of better knowledge. The models are fit via ordinary least squares regression in the first step. The resulting noise is estimated with ARIMA models, which are then used for future simulation to consider the uncertainty in the forecast.

The remaining 1,482 PCs are assumed to be random walk processes and are simulated accordingly. The resulting 10,000 trajectories of the future course of the PCs are transformed back into forecasts of the ASNSNMs through 2040. This process is essentially followed for mortality and fertility forecasting as well. The results of the forecasts are presented in Section 4 among other simulation outcomes.

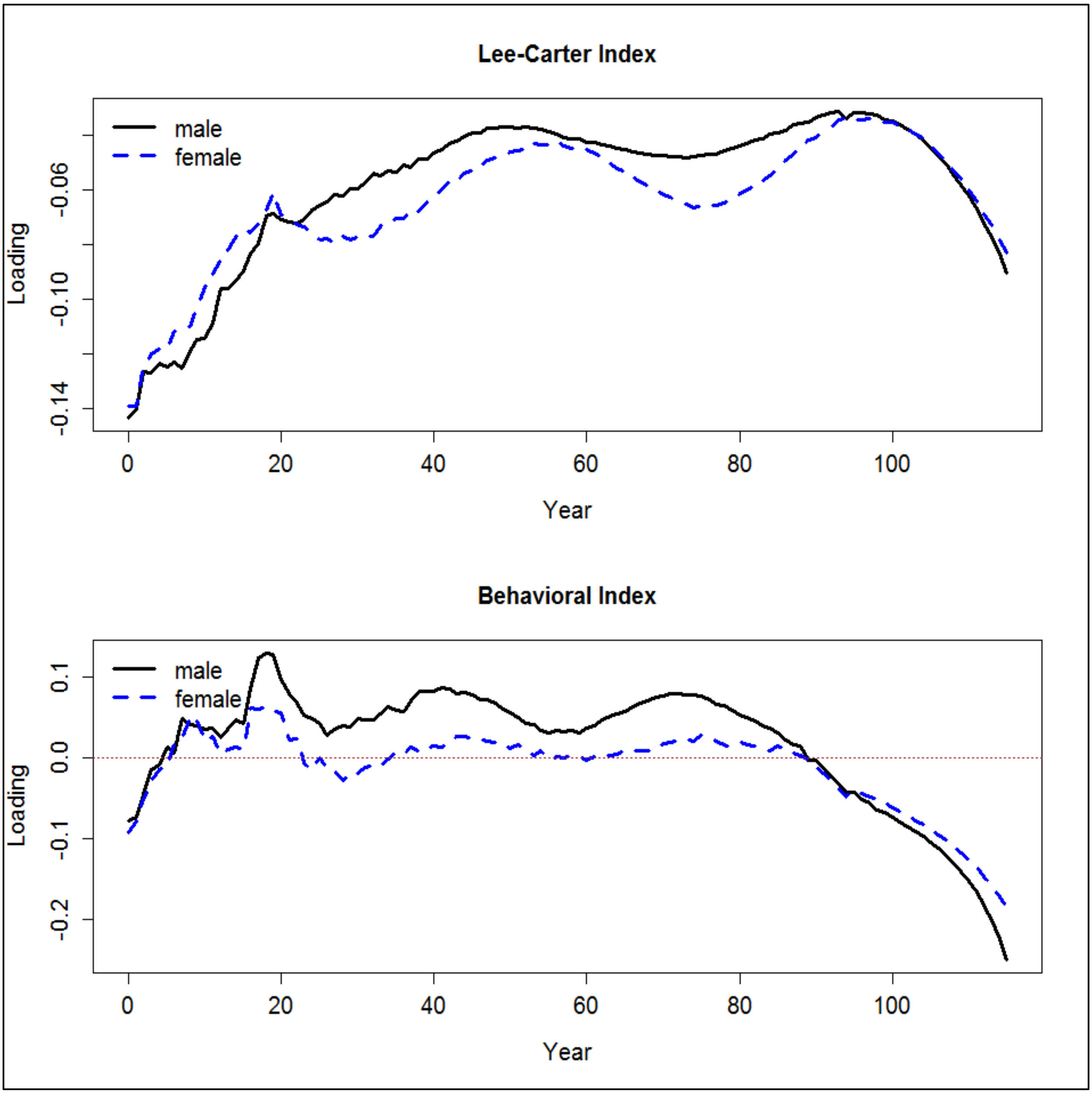
Figure 3. Principal Component Forecast for Migration



Source: Own calculation and design

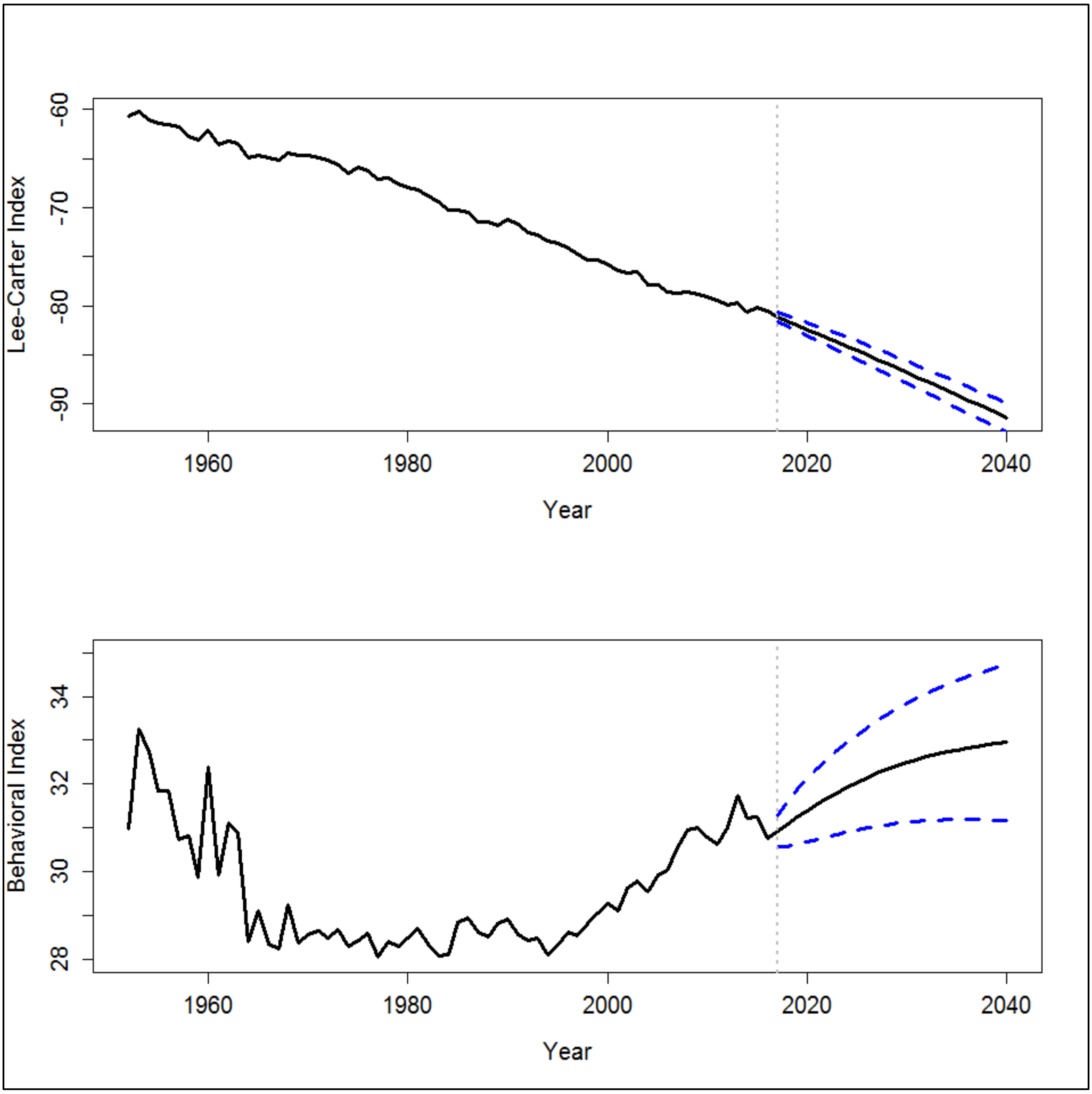
The partial model for mortality is based on Vanella (2017a), where the Destatis data on deaths by sex and age (Destatis 2016b, 2017c, 2018c) as well as the end-of-year population for the years 1952-2016 (Destatis 2015b, 2015c, 2015d, 2016c, 2017d, 2018d) are used to estimate age- and sex-specific mortality rates *(ASSMRs)* for 0-94 year-olds. This procedure has the advantage of deriving adjusted ASSMRs, which include changes in the population due to international migration, directly from our mortality measure. The timing of migration is covered by the ASSMRs, assuming it is similar to the past timing of migration. The ASSMRs for ages 95 and over are estimated by non-linear least squares fitting of logistic models until age 115. Age- and sex-specific survival rates *(ASSSRs)* result from subtracting the corresponding ASSMRs from 1. A PCA is performed for the ASSSRs. Vanella (2017a, 543-8) identified the first two PCs resulting from the PCA as a classical Lee-Carter Mortality Index and a Behavioral Index regarding nutritional and smoking behaviors, respectively, which explain the gender gap in mortality to some extent. The loadings of these two PCs as well as their future forecasts are illustrated in Figures 4 and 5.

Figure 4. Loadings of the Lee-Carter Index and Behavioral Index



Source: Own calculation and design

Figure 5. Forecast of the Lee-Carter Index and Behavioral Index

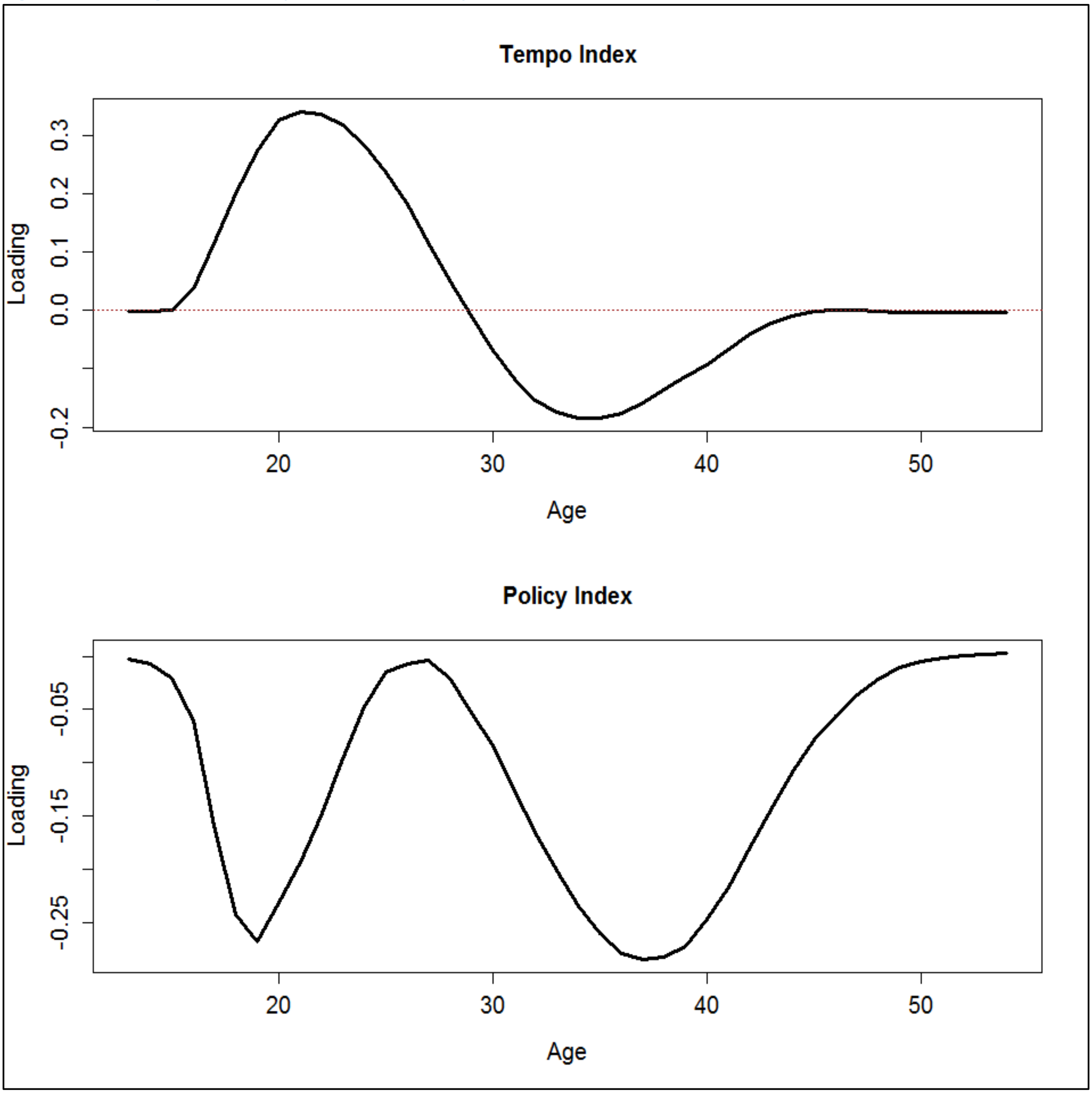


Source: Own calculation and design

The development of the Lee-Carter Index shows a general trend of decreasing mortality over all age groups. The expected increase in the Behavioral Index reflects convergence in nutritional and smoking behaviors between males and females.

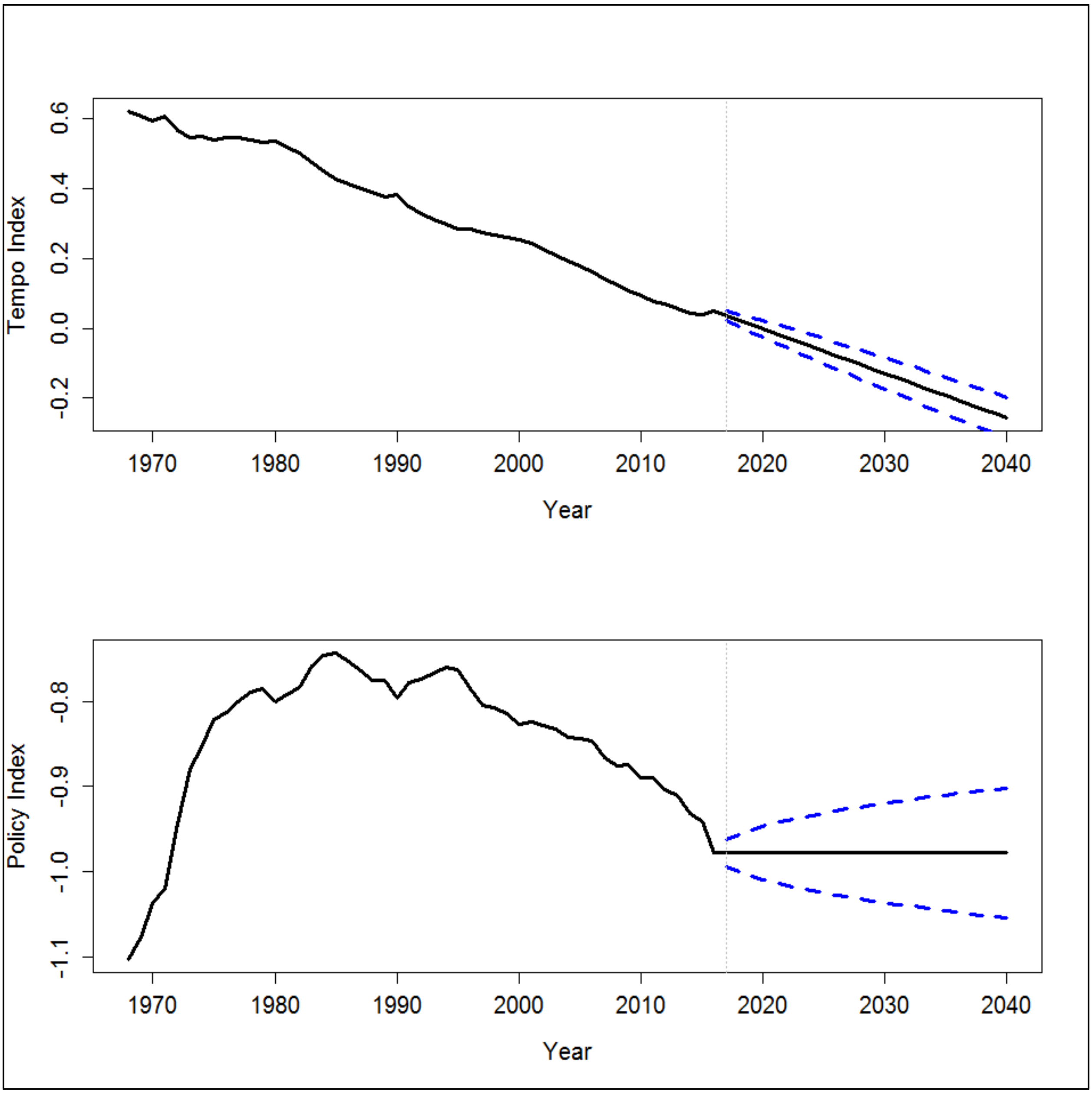
Regarding fertility, we use data on age-specific births among individuals aged 15 to 49 for the years 1968 to 2016 provided by Destatis directly or downloaded from GENESIS-Online (GENESIS-Online Datenbank 2018d, Destatis 2007, 2014a, 2014b, 2018e, 2018f) together with the age-specific data on the female population of reproductive age. Specific birth data on younger or older mothers are not available; therefore, we estimate these data by geometric extrapolation as suggested by Vanella and Deschermeier (2018b, 8-9). We derive age-specific fertility rates (*ASFRs)* by dividing age-specific births by the corresponding mean age-specific female population for the respective year. As proposed by Vanella and Deschermeier (2018b), we run a PCA on the ASFRs for mothers aged 13-54 years for the base period 1968-2016. This time horizon was proposed in that paper because it shows fertility developments after the second wave of the feminist movement. The authors have shown that the first PC represented the tempo effect in fertility, whereas the second PC is associated with the general quantum of fertility and is somehow influenced by family policy. Figure 6 illustrates the loadings of these two PCs.

Figure 6. Loadings of the Tempo Index and Policy Index



Source: Own calculation and design

Figure 7. Forecast of the Tempo Index and Policy Index



Source: Own calculation and design

Figure 7 shows the historical courses of these two variables with the forecast until the year 2040. The forecast of the Policy Index is conditional on the assumption that transfers in family policy are kept constant on the level of 2015, adjusted to inflation.[6]

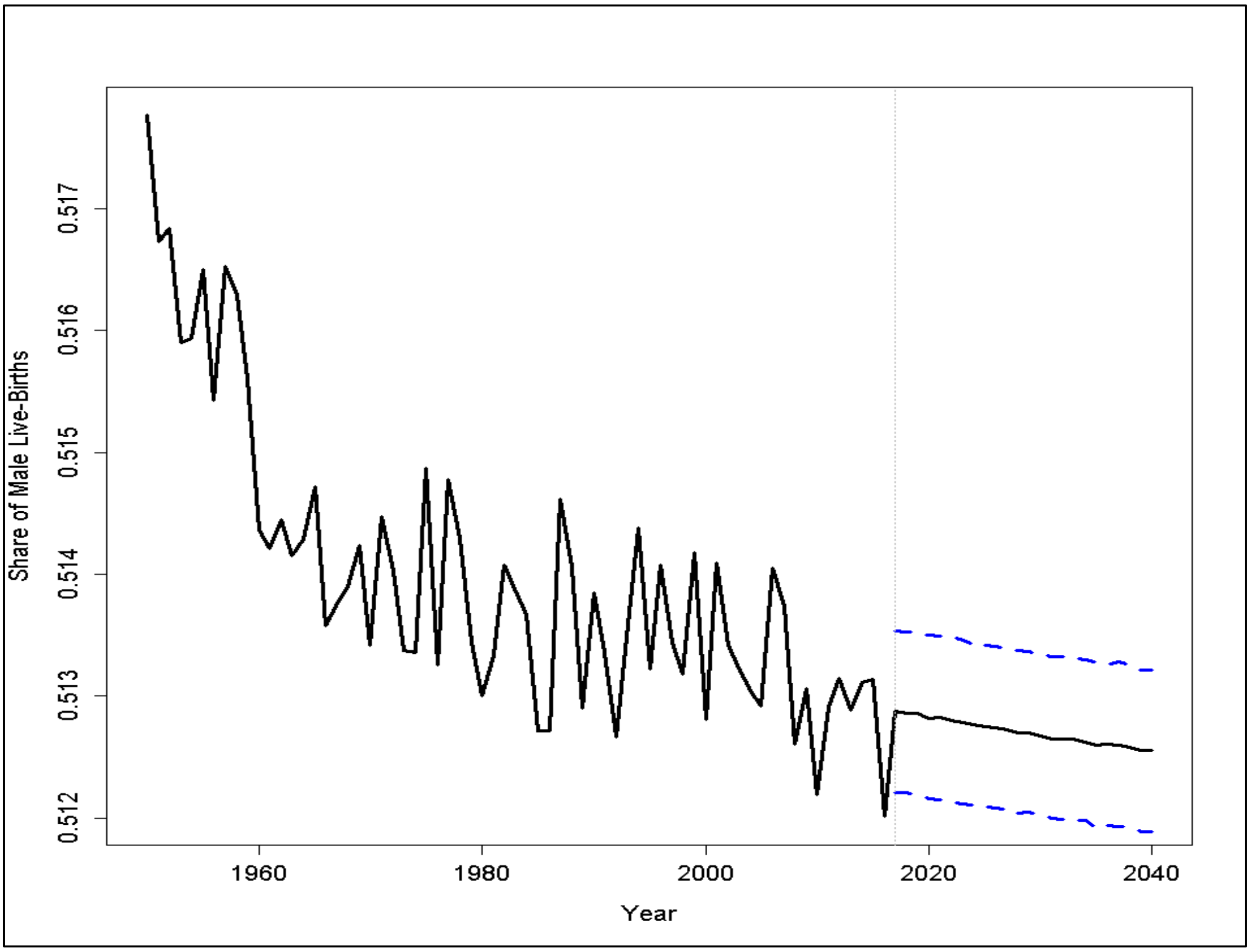
The gender of the children will be simulated after computing the birth numbers. Therefore, we calculate the ratio of males among all live-births annually based on the sex-specific birth numbers in Germany from 1950 to 2016 extracted from GENESIS-Online (GENESIS-Online Datenbank 2018a). We then fit a logistic ARIMA model to the data for simulation of the birth

---

[6] More details on this can be found in Vanella and Deschermeier (2018b, 13-7).

ratio until 2040. The ratio's historical course alongside the median forecast and 75% PIs is given in Figure 8.

Figure 8. Forecast of the Sex Ratio



Sources: GENESIS-Online Datenbank 2018a, Own calculation and design

An apparent trend of a decreasing ratio of male births is evident over the analyzed horizon. This trend can also be observed in other industrialized countries since at least the 1970s (Davis et al. 2007, 941-3, James 2000, 1179-82). Although various studies individually report some evidence that environmental factors such as weather (Helle et al. 2008), exposure to toxins (see e.g. Davis 941-2), and nutritional behavior (Mathews et al. 2008, 1662-6) have some influence on a baby's sex, none of the findings explain the observed trends of decreasing ratios of male births. Considering the clear basic trend since 1950, assuming that the trend will continue over the forecast horizon is plausible.

We now describe the procedure for population forecasting with our model. Let $P_{x,y,g,t}$ denote the population aged x years at the end of year y for sex g in trajectory t. The population update is performed through the following step-wise process.

Step I:

The forecast begins with an adjustment of the base population with regards to international migration flows in the first forecast year *y+1***.** The addition of international net migration aged *x+1* years of sex g during year *y+1* and in trajectory t ($M_{x+1,y+1,g,t}$) to $P_{x,y,g,t}$ leads to the hypothetical subpopulation $\tilde{P}_{x+1,y+1,g,t}$ at the end of year *y+1* without any deaths:

$$\tilde{P}_{x+1,y+1,g,t} = P_{x,y,g,t} + M_{x+1,y+1,g,t}.$$

Step II:

The actual number of survivors from $\tilde{P}_{x+1,y+1,g,t}$ at the end of *y+1* is calculated through multiplication with the adjusted age- and sex-specific survival rate (*ASSSR*) $s_{x+1,y+1,g,t}$ for persons aged *x+1* years of sex *g* in year *y+1* and in trajectory *t*:

$$P_{x+1,y+1,g,t} = \tilde{P}_{x+1,y+1,g,t} * s_{x+1,y+1,g,t}.$$

Step III:

The mean female population in *y+1* in the reproductive age group is approximated:

$$F_{x,y+1,w,t} = \frac{P_{x-1,y,w,t} + P_{x,y+1,w,t}}{2}.$$

Step IV:

The live births $B_{y+1,t}$ are estimated:

$$B_{y+1,t} = \sum_{x=13}^{54} F_{x,y+1,w,t} * f_{x,y+1,t},$$

where $f_{x,y+1,t}$ denotes the ASFR for females aged x years in year *y+1* in trajectory *t*.

Step V:

The shares of the male $(B_{y+1,m,t})$ and female $(B_{y+1,w,t})$ live birth numbers are calculated:

$$B_{y+1,m,t} = B_{y+1,t} * r_{y+1,m,t},$$

with $r_{y+1,m,t}$ representing the ratio of male live births in year *y+1* in trajectory *t*. Subsequently, the female birth numbers are

$$B_{y+1,w,t} = B_{y+1,t} * (1 - r_{y+1,m,t}).$$

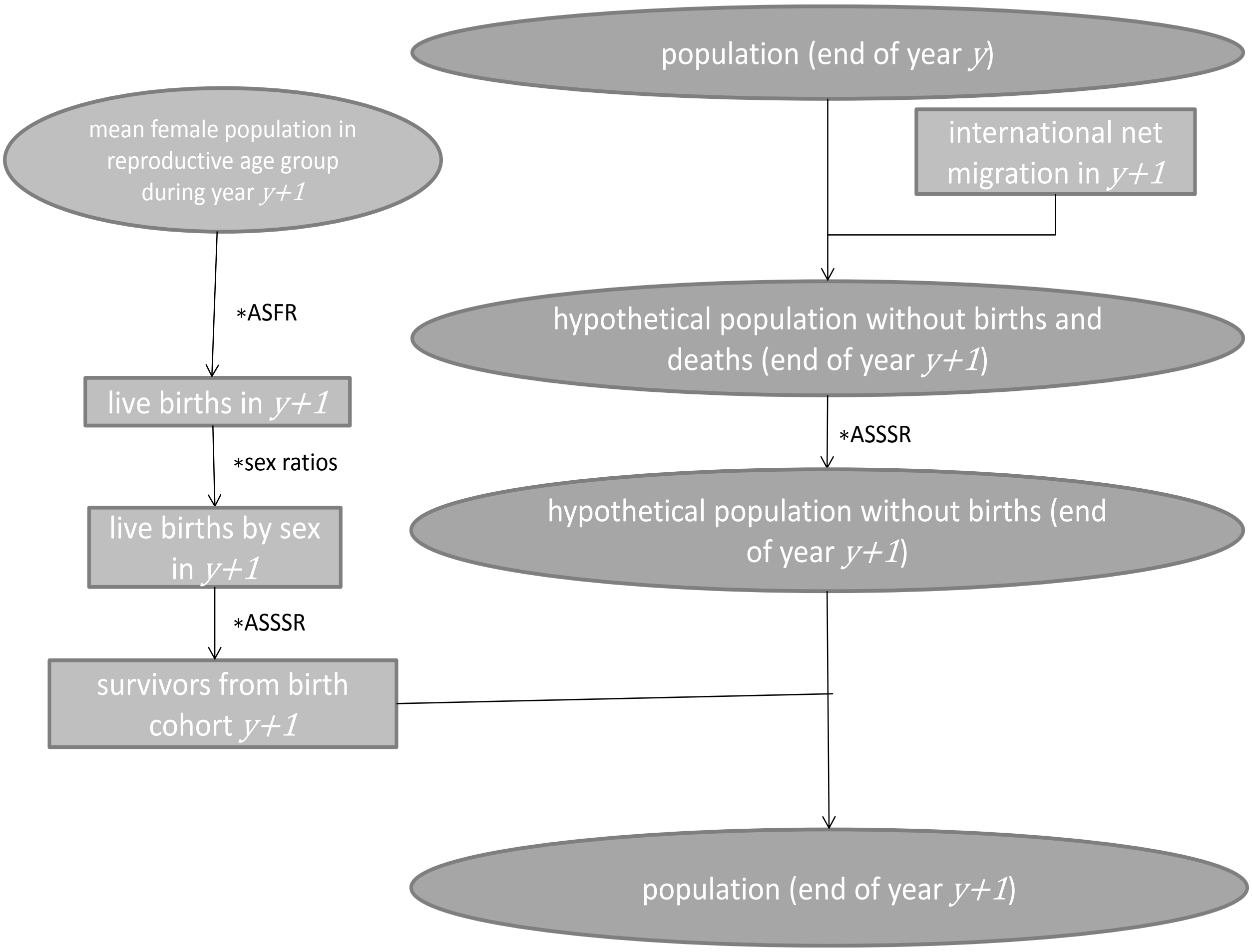
Step VI:

The number of survivors among the children born in $y+1$ is calculated:

$$P_{0,y+1,g,t} = B_{y+1,g,t} * s_{0,y+1,g,t}.$$

In this way, the population by sex and age in year $y+1$ in trajectory t is obtained. This process is then used to stochastically forecast the population by sex and age until the year 2040. The algorithm is illustrated in Figure 9.

Figure 9. Process of Annual Population Update



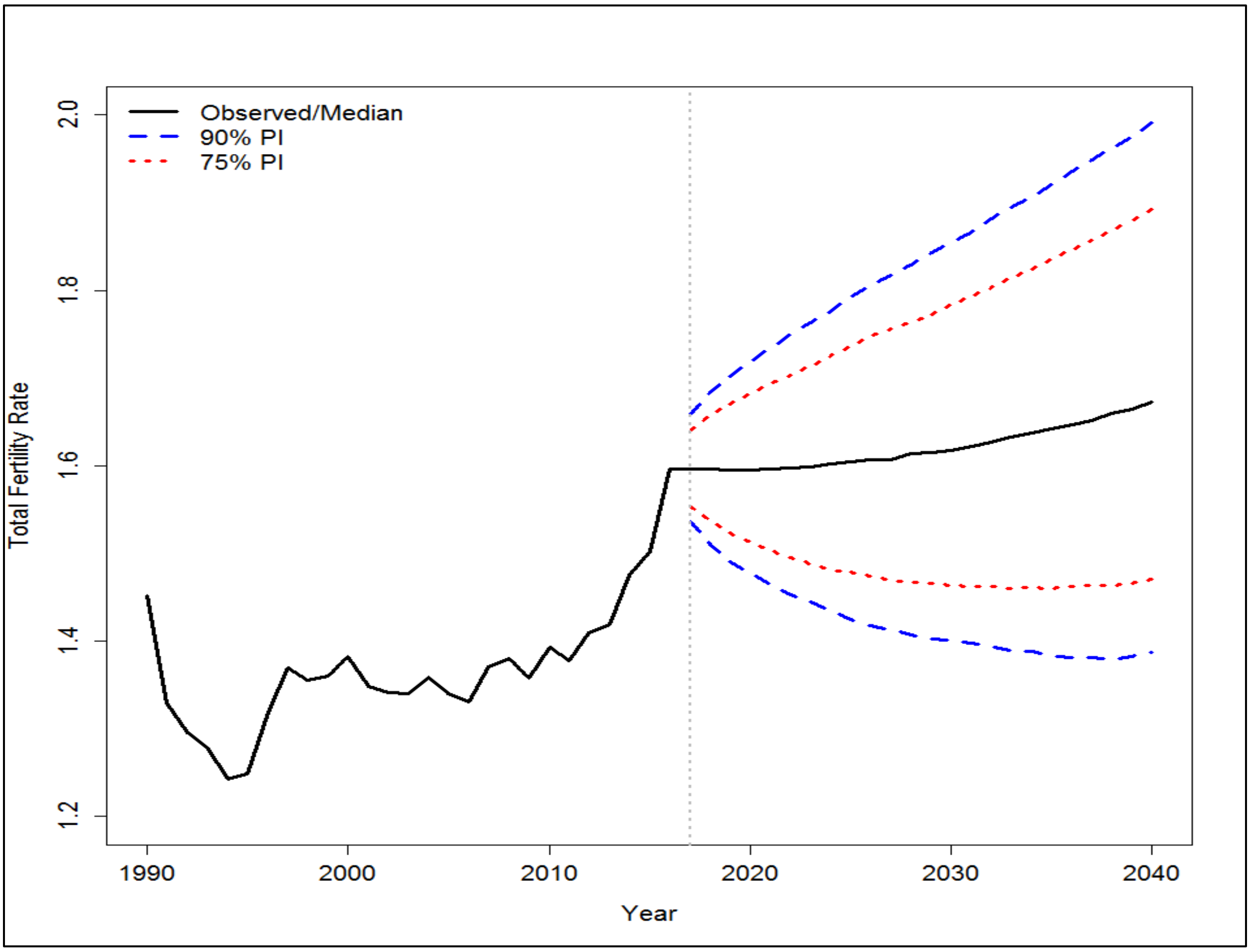
Source: Own design

4 Population Development in Germany until 2040

The combination of the resulting trajectories for the demographic components results in a probabilistic cohort-component model for forecasting the age- and sex-specific population for the ages 0-115 years. The initial population for the forecast is the age- and sex-specific population reported by Destatis for December 31, 2016 (Destatis 2018d). Population numbers for ages 100 years and older are not available in detail but are instead aggregated into an upper age group. Therefore, we estimated the population in this age group through geometric extrapolation until age 115.

In Section 3, we described the partial models for forecasting of the demographic components. Now, we provide a selection of the results from the forecasting procedure. We will keep the overview short since more detailed results for the age-specific measures can be found in previous papers. The fertility model results in 10,000 trajectories for all ASFRs. The total fertility rate *(TFR)* is simply the sum over all ASFRs, thus representing a good indicator for the overall fertility level during a certain period. The median predicted TFR and 75% and 90% PIs are illustrated in Figure 10.

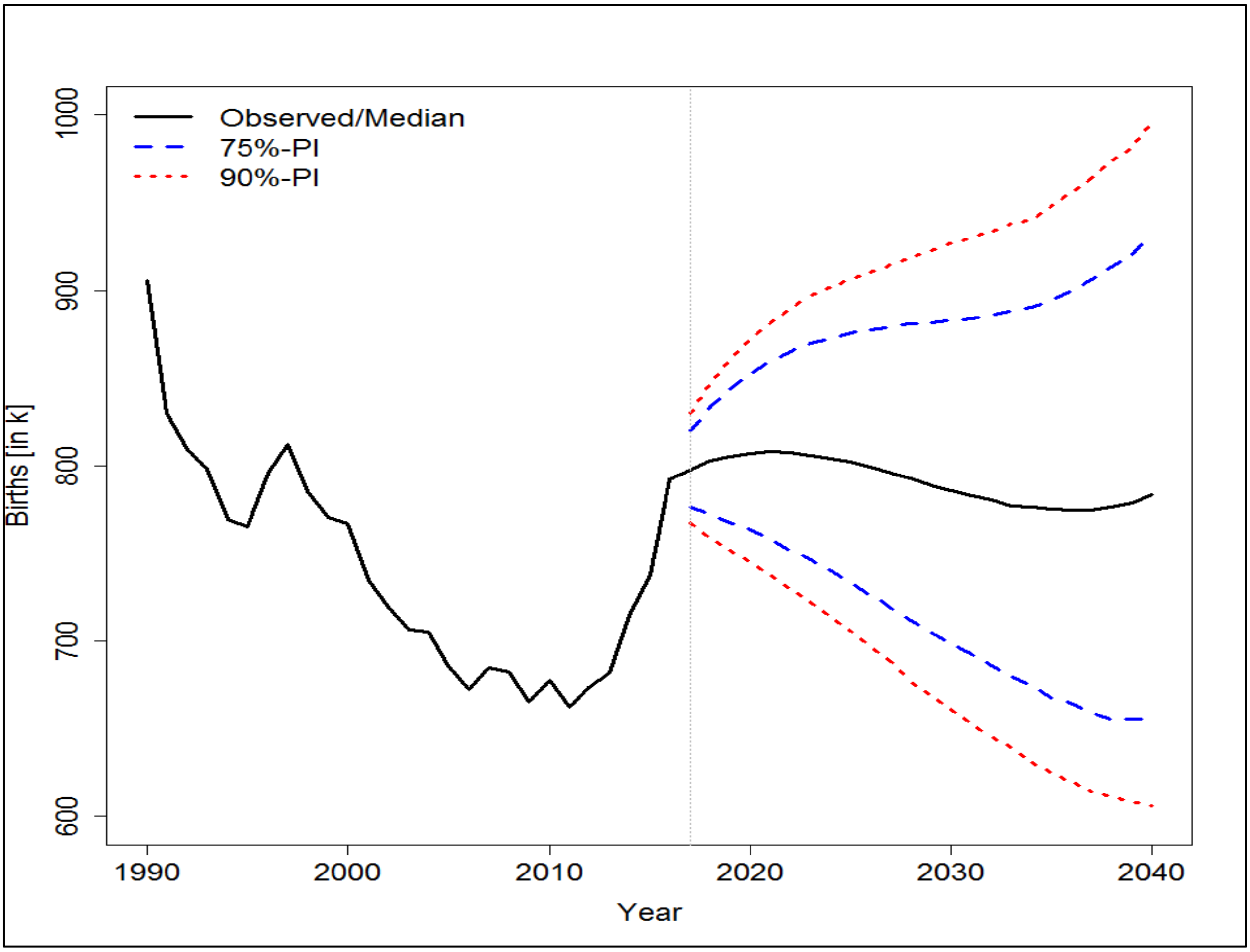
Figure 10. Forecast of the Total Fertility Rate



Sources: GENESIS-Online Datenbank 2018d, Destatis 2007, 2014a, 2014b, 2018e, 2018f, Own calculation and design

Considering the probable absence of any landslide innovations in family policy in the future, we can expect a further but moderate increase over the forecast horizon. The median of the forecast is 1.67 and lies within the 75% PI of [1.47; 1.89] and the 90% PI of [1.39; 1.99]. By multiplication of the ASFRs with the corresponding female population, the birth forecast is completed. The results are given in Figure 11.

Figure 11. Forecast of Birth Numbers

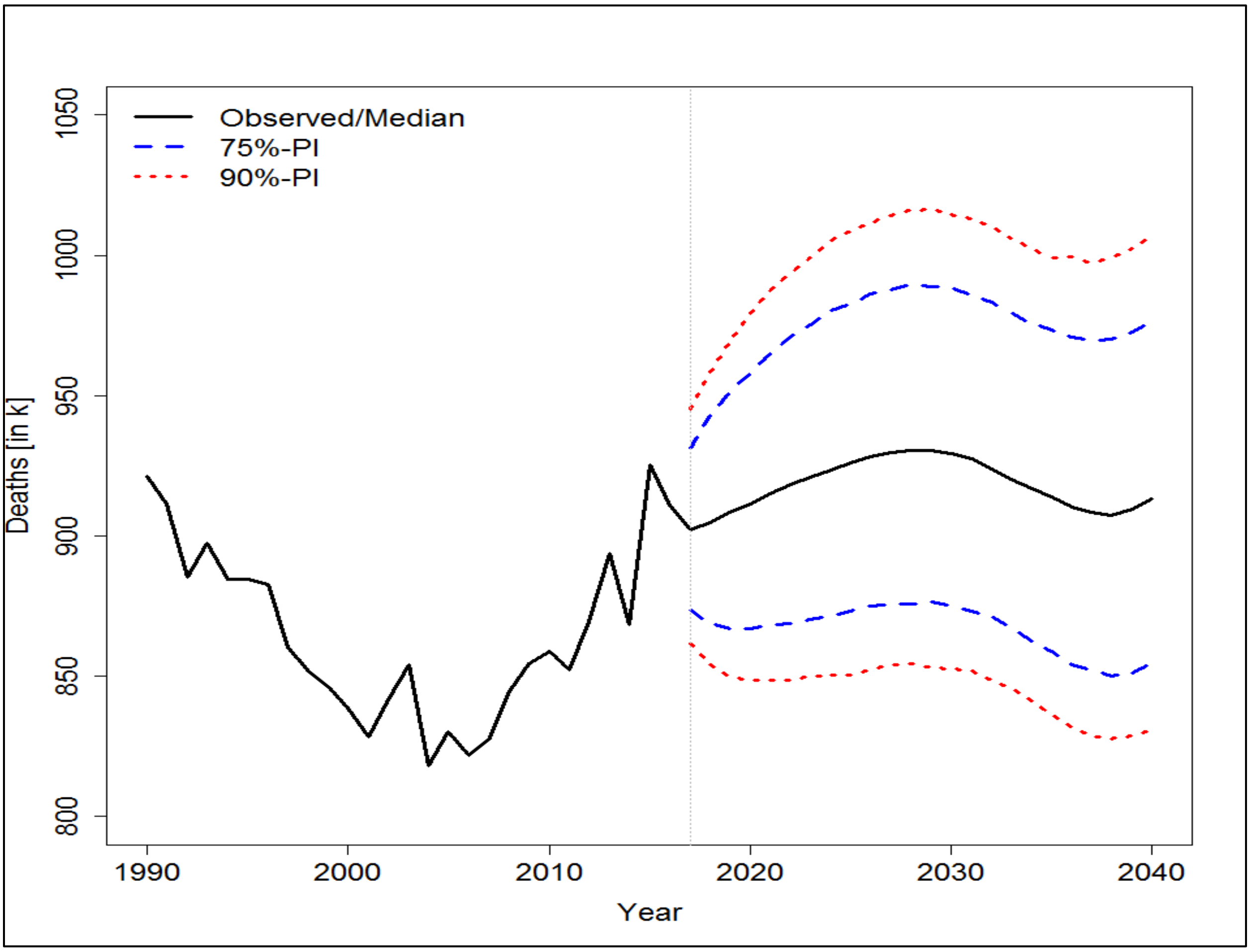


Sources: GENESIS-Online Datenbank 2018a, Own calculation and design

The increasing trend in births, as witnessed since 2012, are expected to continue until 2020. Regardless of the increasing TFR, the birth numbers will probably subsequently decrease moderately because most children are born by mothers over 29 years of age, as shown by Vanella and Deschermeier (2018b, 19). This decrease can therefore be explained by the decreasing number of births at the beginning of the 1990s, as shown at the left-hand side of the graph. The median increase during the second half of the 2030s stems from the increasing TFR together with almost stagnating birth numbers during the cohorts 2005 and 2011, which by then will be in their reproductive phase.

Similarly, the death numbers are derived from the ASSSRs and the population update. As shown in Figure 9, deaths can be derived by simulating the hypothetical age- and sex-specific population at the end of some period in some trajectory without deaths and then multiplying this number with the respective adjusted ASSMR to derive the actual number of deaths among this group. The resulting death numbers are illustrated in Figure 12.

Figure 12. Forecast of Deaths

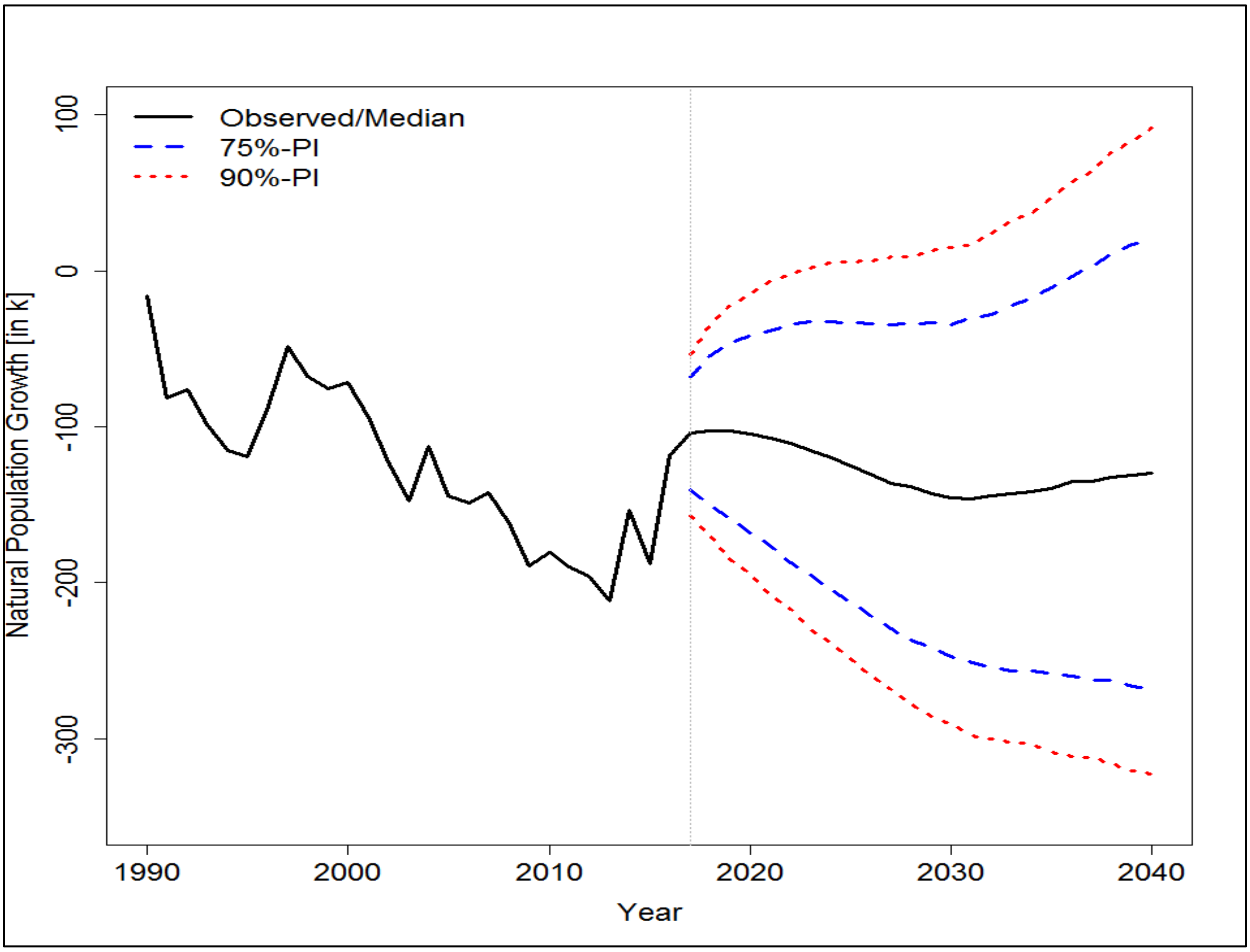


Sources: GENESIS-Online Datenbank 2018b, Own calculation and design

A further increase in deaths, as witnessed since the mid-2000s, is probable until the late 2020s. This results from the strong birth cohorts of the 1960s and early 1970s, who until then are all over 60 years of age, about the age group, where mortality risks start increasing strongly. The numbers are expected to decrease a bit until the late 2030s, after which on average another increase will occur. At that point, many of the immigrants coming from Germany since reunification will be in their 60s and older, therefore witnessing higher mortality risks themselves.

By subtracting the death numbers from the birth numbers, we calculate the natural population growth, whose forecast can be derived indirectly from the birth and death forecasts as well. The results of the natural population growth forecast is illustrated in Figure 13. A slight negative tendency is probable. At high likelihood, the deaths will exceed the births over the forecast horizon. The simulation study gives a probability of just 15.9% in 2040 for the birth numbers to exceed the death numbers.

Figure 13. Forecast of Natural Population Growth



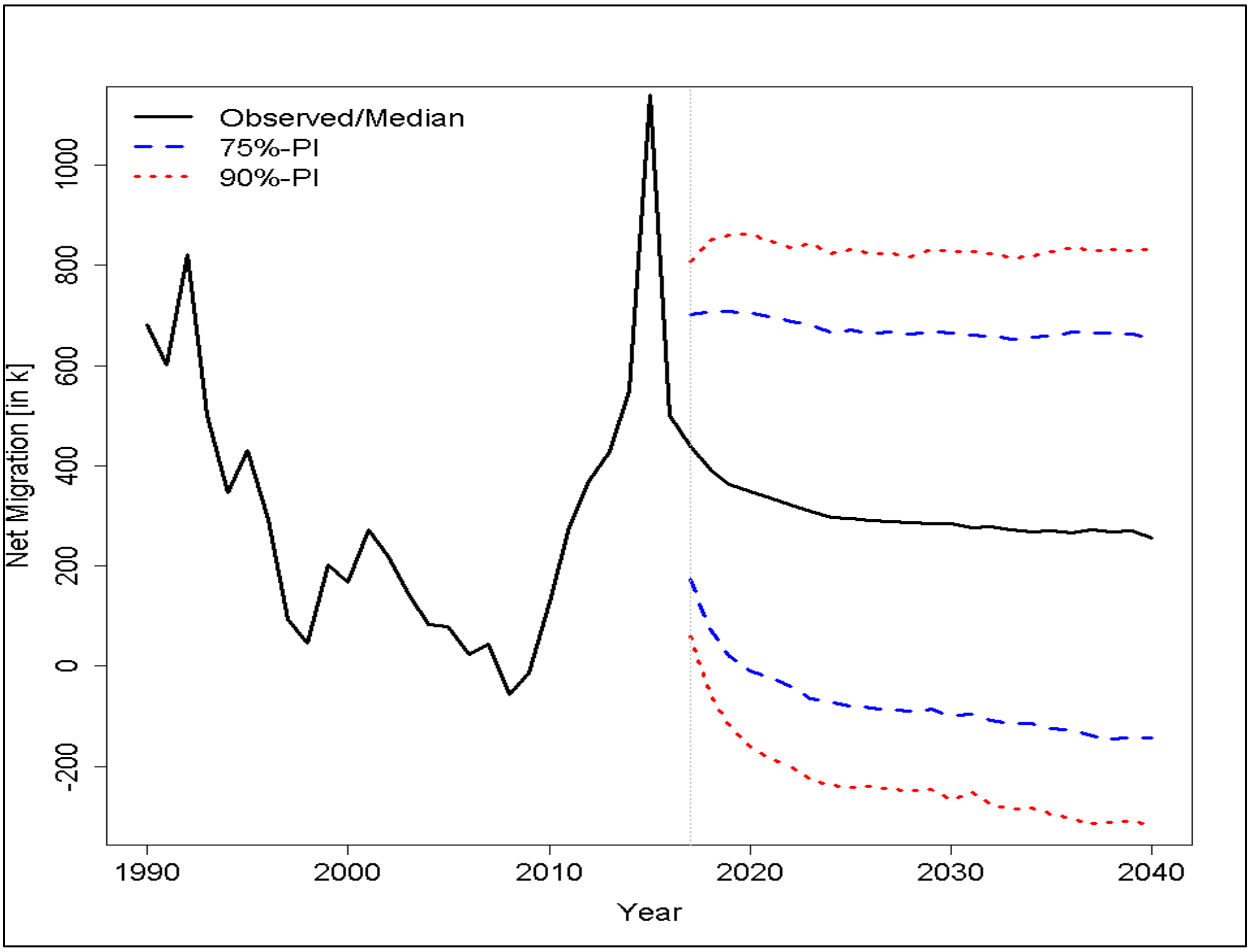
Sources: GENESIS-Online Datenbank 2018a, GENESIS-Online Datenbank 2018b, Own calculation and design

Counterbalancing the shrinking population due to natural population decrease is the international net migration. The forecast method for the ASNSNM numbers has been explained in Section 3, the results of the simulation are cumulated into the total net migration for illustration purposes in Figure 14.[7]

---

[7] More detailed results, although based on the jump-off year 2015, can be found in Vanella and Deschermeier 2018a, 274-6.

Figure 143. Forecast of Net Migration



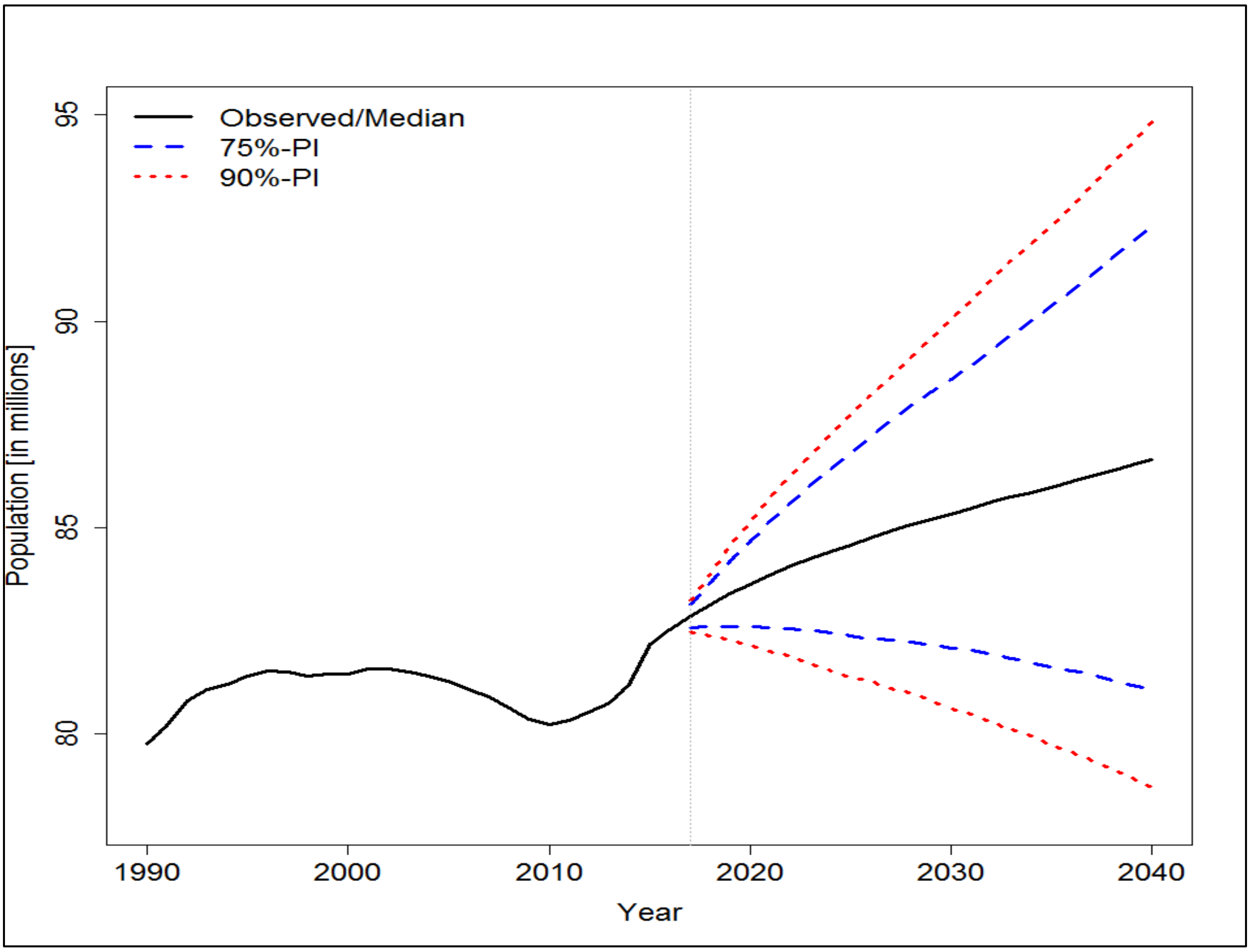
Sources: GENESIS-Online Datenbank 2018e, Own calculation and design

The median scenario gives a slightly decreasing net migration, whereas some cyclic course due to economic cycle is probable. In general, the high uncertainty in migration forecasting is obvious, but in general a positive net migration is very likely. The median of net migration in 2040 is 255,334 persons. This is a higher balance then most previous forecasts provide, that were calculated before the record influx of 2015. As many bigger cities of origin of the refugees, especially in Syria, are mostly devastated by war (McKenzie 2018, Pleitgen 2017), it seems unlikely that there will be a mass emigration out of Germany in the years to come, as one might expect due to experience from past refugee crises. Furthermore the results reflect the strong past development of the economy in Germany. This trend is probable to remain stable in the future (OECD 2017, 130-3). The attractive labor market is likely to attract more people in the future (Fuchs et al. 2018, 49-54), especially within the EU due to the unrestricted free movement of workers (Vanella and Deschermeier 2018a, 274-7). Total net migration in 2040 is estimated above zero at 76.77% probability.

The high importance of positive net migration, especially in the younger ages, shall be mentioned to fill the shortages occurring in the labor market due to overaging. We stress that the effect of migration on the labor market and the social security system very much depends on the skill level and education of the immigrants. Especially in cases of refugee migration, where education in many cases either is at a relatively low level or is not accepted by German standard, it usually takes a long time for the immigrants to fully integrate into the labor market (Brücker et al. 2017b).

The algorithm explained in Figure 9 gives us the forecast of the end-of-year population for each year over the forecast horizon, both sexes and each age from 0 to 115 years. We will finally present some of the results at this point. Figure 15 shows the forecast of total population until the year 2040 with 75% and 90% PIs. In contrast to many earlier studies on Germany (see Section 2), the population is expected to increase moderately over the forecast horizon due to high, yet decreasing, net migration, an increasing TFR and decreasing mortality. Contrary to common belief and based on our findings, there is no empirical evidence for a decrease in population size through the chosen forecast horizon. Although realistic, the likelihood of the decrease is relatively small.

Figure 15. Forecast of Population Size



Sources: GENESIS-Online Datenbank 2018f, Destatis 2016d, Own calculation and design

Table 1 shows the forecast and projection results for Germany for a selection of the studies mentioned in Section 2. The percentiles were chosen as given to offer the opportunity to compare our results to those of other studies without biased results due to different quantiles. Studies not mentioned here, like Lipps and Betz (2005) or Deschermeier (2016) didn't provide the corresponding percentiles or had a shorter forecast horizon.

Table 1. Approximate Forecasts and Projections for the Population of Germany in 2040

Study	5th Percentile	Median	95th Percentile
Bomsdorf et al. (2008)	**68**	**76**	**83**
Härdle and Myšičková (2009)	**76**	**79**	**83**
Dudel (2014)	**74**	**79**	**84**
UN (2017)	**78**	**81**	**84**
own forecast	**79**	**87**	**95**

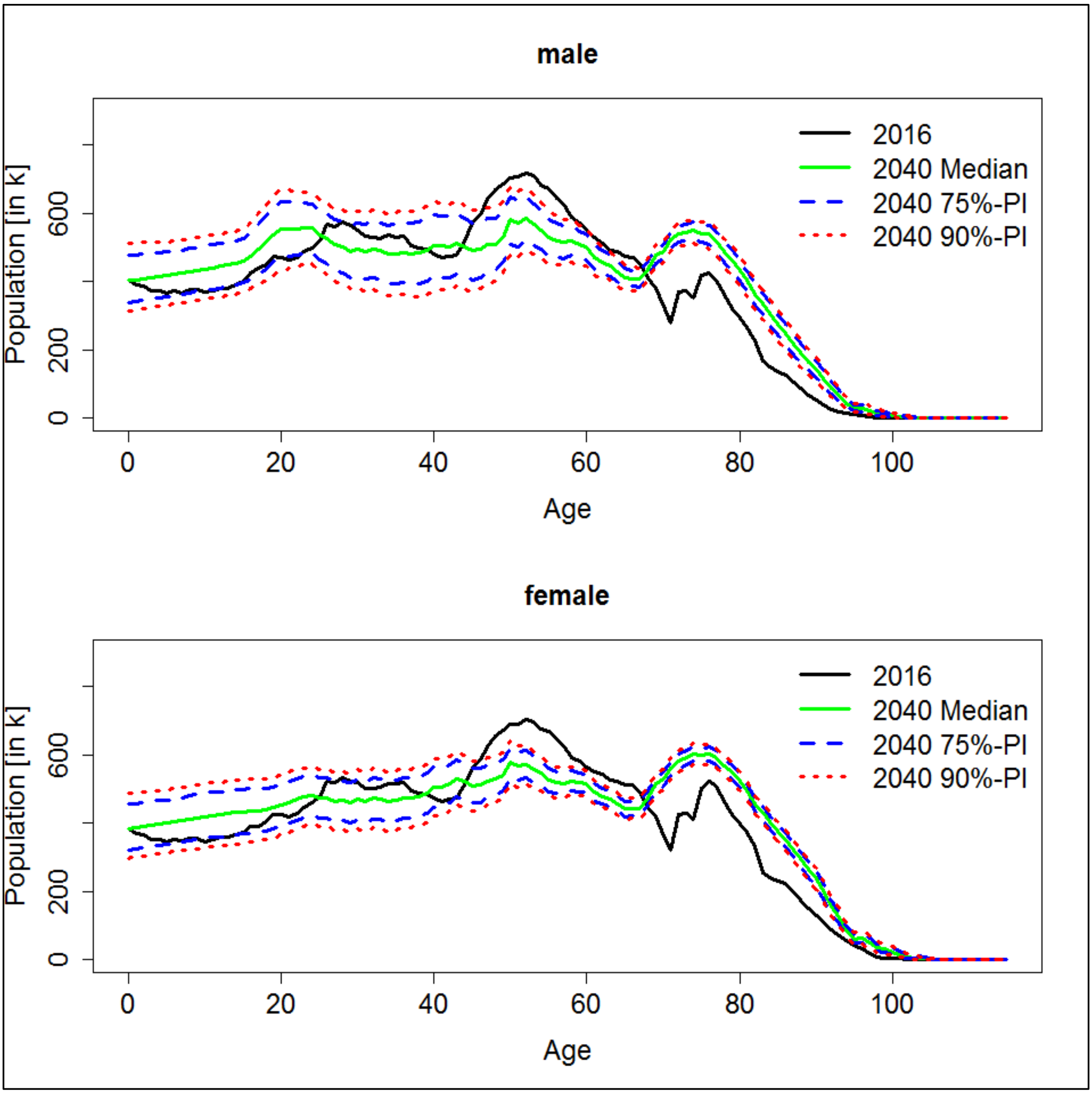
Source: Own calculation and design

Our median forecast of the total population is substantially above those of the other studies. By contrast, the other studies mainly assume a constant TFR and a smaller magnitude of migration. Vanella (2017b, 25-6) showed that naïve fertility forecasts (which assume a constant TFR with a fixed age schedule) empirically perform rather poorly. Moreover, the presented studies are outdated in the sense that they were conducted before the refugee crisis since 2015 and the above average net migration since 2010, caused by the European debt crisis. These developments mark significant structural long-run changes in migration. At the time most of these studies could not be foreseen. Bujard and Dreschmitt (2016, 337-42) stress this point and argue, based on historical migration data and qualitative comparison with the refugee crisis during the early 1990s, that migration levels higher than those assumed in previous studies are realistic.

Citing the authors, given the attractiveness of Germany for migrants and politically induced migration, a long-term net migration of 300,000 seems a real possibility. They showed that this level of migration would lead to an increase in total population until the early 2030s. Even if 300,000 long-term net migration is certainly a bold claim, our results quantitatively show, that a quarter million net migration in the long term is very probable. Furthermore, our study shows that the earlier studies underestimate the uncertainty in the forecast with 90% PIs of 6-15 million persons. By contrast, according to our forecast, the ceteris paribus population in 2040 will be between 78.713 and 94.829 million people at a 90% probability level, with a median outcome of 86.647 million. The actual estimate for the total population on December 31, 2017 is approximately 82.792 million people (GENESIS-Online Datenbank 2018f), which lies within the 20%-PI quantified with our model. The absolute difference to our median estimate of 82.851 million is quite small as well, showing a good fit of our model, at least for the first year of the forecast.

In many cases (like in social security), the structure of the population is of higher importance than its size per se. Therefore, Figure 16 gives an overview of the age structure of the population in 2016 compared to the forecast in 2040 with PIs for both sexes.

Figure 16. Population by Sex and Age in 2016 and 2040

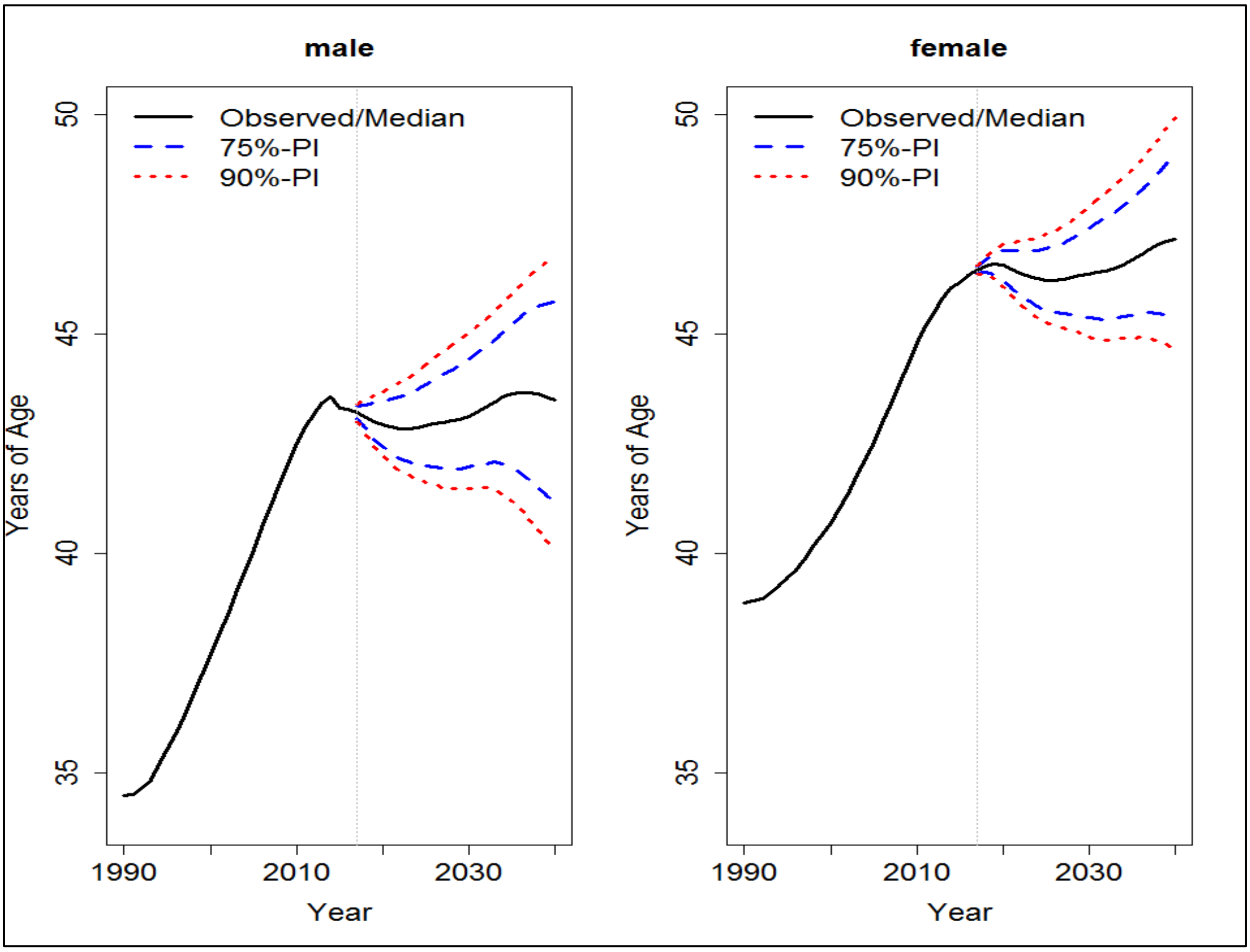


Sources: Destatis 2018d, Own calculation and design

We observe the 25-year shift in the population. In general, there is greater uncertainty for males. Whereas the retirement-age population can be predicted relatively well, the uncertainty in the future working-age population is rather large for males due to the higher uncertainty in the migration forecast for males relative to females (Vanella, Deschermeier 2018a, 274-7). The uncertainty in the population of persons under 25 years of age mostly arises from the fact that this portion of the population has not been born yet but also to the relatively high uncertainty in international migration.

The aging of the population is of high social and economic importance. Therefore, in addition to the overall age structure, the median age of the male and female populations is considered as a summary indicator for the future age schedule of the population. The median age of the population can be obtained from the simulation results because it is the exact age that cuts the population in half. This computation for all 10,000 trajectories can be used to extract PIs for the median age, similar to the computation of median life span conducted by Vanella (2017a: 548-52). The results of this analysis are shown for both sexes in Figure 17.

Figure 17. Median Age by Sex until 2040



Sources: Destatis 2015b, 2015c, 2015d, 2016c, 2017d, 2018d, Own calculation and design

We observe a rejuvenation effect for the upcoming years due to high net migration during this period, as illustrated in Figure 14, and to increasing birth numbers, as shown in Figure 11. The high net migration around the year 2015 combined with the high forecast values for the upcoming years leave a mark in the age structure of Germany. This can be seen in the age structure for the male and female populations in the year 2040. By that time, the majority of the population that immigrated during the high influx phase will be approximately 50 years old, while the baby boomer generation will be in their seventh decade of life. Over the forecast horizon, the median age traces this development by a rejuvenation effect for men and women. The probable decrease in the number of births after the early 2020s and decreasing net migration and mortality (Vanella 2017a, 550) lead to an aging of the population structure, as represented by the increasing median ages after that point. Since a larger portion of migrants is male (Vanella and Deschermeier 2018a, 274-6), the rejuvenation is stronger for males than for females.

When the proportion of the male population on December 31, 2011 in Germany has been about 48.8%, and already increased to 49.3% on December 31, 2016, it is expected to so slightly increase to 49.4% in the median trajectory on December 31, 2040. This can be attributed to a small degree to the slowly converging mortality rates of males towards those of females (Vanella 2017a, 551), but mainly to the higher male net migration by non-European citizens into Germany (Vanella and Deschermeier 2018a: 276), which is counterbalanced by decreasing likelihood for bearing a male child (see Figure 8) and the still higher female life span (Vanella 2017a, 551). During the late 2030s, the male population is expected to again become slightly younger, which results from the higher birth numbers altogether with more deaths, as shown in Figures 11 and 12.

As we have shown by some important measures, our model provides a wide range of detailed analyses targeting specific topics of interest. The forecast results offer the possibility for a wide range of future studies, e.g., analyzing the effects of population changes on social security, the labor market or housing demand.

5 Conclusions, Limitations and Outlook

This paper proposed a probabilistic cohort-component approach for population forecasting by sex and age. It was applied to predict the population of Germany until the year 2040. Germany witnessed a record migration influx in 2015 due to the refugee movement, especially from Syria, Iraq and Afghanistan, in combination with the challenging economic situation in many countries in Southern and Eastern Europe. The record net migration marks a considerable event for Germany’s demographic development. The expected long-term decrease in the population does not appear to hold based on our findings. The results provide essential data on the consequences of the current trends for decision makers, planners and scientists.

The model predicts the population by age and sex of Germany until the year 2040. The forecast is conducted as a composite of three time series models based on PCA for the three demographic components fertility, international migration and mortality by sex and age. The fertility model is conditional on political intervention as well, considering reforms in family policy to some extent. The method is specified for Germany, but it can be applied to other countries or regional units, for which sufficiently long time series data for the demographic components are available. Stochastic modeling of the population produced point estimates of the future population in addition to a measure of the future uncertainty via prediction intervals. The results may be disaggregated or aggregated almost arbitrarily regarding sex, age and level of uncertainty.

The results indicate a high probability of a larger future population for Germany compared to the results of other studies. Our study shows that the increase in the population is mainly due to a larger amount of the older population. The uncertainty in the forecast mainly arises from fertility and migration, since mortality can be forecast with low uncertainty, due to a stable development in the past.

The model is well suited for regular updating and does not require large amounts of data input since it is restricted to demographic variables and uses official statistics provided by Destatis. One interesting innovation is the detailed reporting and probabilistic quantification of the disaggregated population for all ages and both sexes; therefore, the results offer many possibilities for future forecast studies that require disaggregated population data as inputs, e.g., research on social security or life insurance.

Our method is restricted to quantitative methods; therefore, past unobserved trends are not considered in the future. Nevertheless, for all demographic variables, the input data span at least as long of a time horizon as is forecast; thus, we believe that all realistic trends that might be observed during the time horizon are included in the model. The addition of expert knowledge would be possible, if the forecaster thinks the past trends insufficiently cover the possible future outcome. The model suffers from a small input time horizon because the migration data are restricted back to the years 1990. Older data is not representative because of then and the overall very different geopolitical situation in Eurasia back then. Furthermore, fertility is difficult to forecast, since it is strongly influenced by policy as well. We tried to induce this effect to some extent into the model as well, following a ceteris paribus assumption in family policy to avoid bias as well as possible. Our forecast horizon is 2040 and not 2060 or 2100, as in other studies, since we do not intend to create misinterpretations for the far future, for which forecasts are not possible with the available data.

A larger forecast period would be interesting but cannot be achieved via responsible statistical modeling. Thus, the future availability of input data suited for model estimation will improve the quality of our models and allow for longer forecast horizons. Even with a forecast horizon that reaches only until 2040 the uncertainty is rather large. Most of the risk stems from the uncertainty about future net migration. Empirical updating might be required if the development in the upcoming years differs from our forecast due to political or economic developments. Those structural breaks are not implemented in our simulation approach. Furthermore a more detailed migration forecast between emigration and immigration might help to better understand the sensitivity of population development.

Acknowledgements

We express our thanks to our colleague Johann Fuchs from the Institute for Employment Research (IAB) in Nürnberg for his very valuable comments on an earlier version of this paper.

References

Alho, Juha M. and Bruce D. Spencer. 2005. *Statistical Demography and Forecasting.* New York: Springer Science+Business Media.

Alkema, Leontine, Adrian E. Raftery, Patrick Gerland, Samuel J. Clark, François Pelletier, Thomas Büttner, and Gerhard K. Heilig. 2011. “Probabilistic Projections of the Total Fertility Rate for All Countries.” *Demography* 48 (3): 815-39.

Aust, Stefan, Manuel Bewarder, Wolfgang Büscher, Martin Lutz, and Claus C. Malzahn. 2015. “Herbst der Kanzlerin. Geschichte eines Staatsversagens.“ *Welt,* November 09, 2015. https://www.welt.de/politik/deutschland/article148588383/Herbst-der-Kanzlerin-Geschichte-eines-Staatsversagens.html.

Bell, William and Brian Monsell. 1991. “Using Principal Components in Time Series Modeling and Forecasting of Age-Specific Mortality Rates.” In *Proceedings of the Social Statistics Section 1991,* edited by American Statistical Association, 154-9. Alexandria: American Statistical Association.

Bewarder, Manuel and Marcel Leubecher. 2016. “Syrische Staatsbürgerschaft wird massenhaft vorgetäuscht.“ *Welt*, June 23, 2016. https://www.welt.de/politik/deutschland/article156496638/Syrische-Staatsbuergerschaft-wird-massenhaft-vorgetaeuscht.html.

Bohk, Christina. 2012. *Ein probabilistisches Bevölkerungsprognosemodell: Entwicklung und Anwendung für Deutschland.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Börsch-Supan, Axel H. and Christina B. Wilke. 2009. “Zur mittel- und langfristigen Entwicklung der Erwerbstätigkeit in Deutschland.“ *Zeitschrift für ArbeitsmarktForschung (ZAF)* 42 (1): 29-48.

Bomsdorf, Eckart, Berhard Babel, and Rafael Schmidt. 2008. “Zur Entwicklung der Bevölkerung, der Anzahl der Schüler, der Studienanfänger und der Pflegebedürftigen: Stochastische Modellrechnungen für Deutschland bis 2050.“ *Sozialer Fortschritt* 57 (5): 125-32.

Booth, Heather. 2006. “Demographic forecasting: 1980 to 2005 in review.” *International Journal of Forecasting* 22 (3): 547-81.

Booth, Heather, Rob J. Hyndman, Leonie Tickle, and Piet de Jong. 2006. “Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions.” *Demographic Research* 15 (9): 289-310.

Bozik, James E. and William R. Bell. 1987. “Forecasting Age specific Fertility using Principal Components.” *Bureau of the Census Statistical Research Division Report Series* 87/19. Washington, D.C.: Bureau of the Census, Statistical Research Division.

Brücker, Herbert, Andreas Hauptmann, and Steffen Sirries. 2017a. *Zuzüge nach Deutschland.* Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.

Brücker, Herbert, Andreas Hauptmann, and Steffen Sirries. 2017b. *Arbeitsmarktintegration von Geflüchteten in Deutschland: Der Stand zum Jahresbeginn 2017*. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.

Bujard, Martin and Kai Dreschmitt. 2016. “Szenarien der Bevölkerungsentwicklung bis 2060: Wie beeinflussen Migration und Geburten Deutschlands Zukunft?“ *Gesellschaft. Wirtschaft. Politik* 65 (3): 333-45.

Bundesamt für Migration und Flüchtlinge. 2016. *Das Bundesamt in Zahlen 2015: Asyl, Migration und Integration.* Nürnberg: Bundesamt für Migration und Flüchtlinge.

Bundesministerium des Innern. 2011. *Migrationsbericht 2009.* Berlin: Bundesministerium des Innern.

Bundesministerium des Innern. 2015. *Migrationsbericht 2013.* Berlin: Bundesministerium des Innern.

Bundesministerium des Innern. 2017. *Migrationsbericht 2015.* Berlin: Bundesministerium des Innern.

Cannan, Edwin. 1895. “The Probability of a Cessation of the Growth of Population in England and Wales During the Next Century.” *The Economic Journal* 5 (20): 505-15.

Census Bureau of England and Wales. 1863. “General Report; with appendix of tables.” A vision of Britain through Time. Accessed November 4, 2017. http://www.visionofbritain.org.uk/census/report/EW1861GEN.

Council on Foreign Relations. 2017. “Islamist Extremism After the Arab Spring.” Interview by Jonathan Masters with Ali Soufan. September 14, 2012. https://www.cfr.org/interview/islamist-extremism-after-arab-spring.

Davis, Devra L., Pamela Webster, Hillary Stainthorpe, Janice Chilton, Lovell Jones, and Rikuo Doi. 2007. “Declines in Sex Ratio at Birth and Fetal Deaths in Japan, and in U.S. Whites but Not African Americans.” *Environmental Health Perspectives* 115 (6): 941-6.

Deschermeier, Philipp. 2011. “Die Bevölkerungsentwicklung der Metropolregion Rhein-Neckar: Eine stochastische Bevölkerungsprognose auf Basis des Paradigmas funktionaler Daten.“ *Comparative Population Studies* 36 (4): 769-806.

Deschermeier, Philipp. 2015. “Die Entwicklung der Bevölkerung Deutschlands bis 2030 – ein Methodenvergleich.“ *IW Trends - Vierteljahresschrift zur empirischen Wirtschaftsforschung* 42 (2): 97-111.

Deschermeier, Philipp. 2016. “Einfluss der Zuwanderung auf die demografische Entwicklung in Deutschland.“ *IW Trends - Vierteljahresschrift zur empirischen Wirtschaftsforschung* 43 (2): 21-38.

Destatis. 2007. *Lebendgeborene nach dem Alter der Mutter 1990: Neue Länder und Berlin-Ost.* Data provided on February 13, 2018.

Destatis. 2009. *Bevölkerung Deutschlands bis 2060: 12. koordinierte Bevölkerungsvorausberechnung – Annahmen und Ergebnisse.* Wiesbaden: Destatis.

Destatis. 2014a. *Lebendgeborene nach dem Alter der Mütter 1968-1990: Alte Bundesländer.* Data provided on June 18, 2015.

Destatis. 2014b. *Lebendgeborene nach dem Alter der Mütter 1991-2013: Deutschland*. Data provided on June 18, 2015.

Destatis. 2015a. *Wanderungen zwischen Deutschland und dem Ausland 1991-2013 nach Einzelaltersjahren und Geschlecht.* Data provided on September 02, 2015.

Destatis. 2015b. *Bevölkerung und Erwerbstätigkeit 2012: Vorläufige Ergebnisse der Bevölkerungsfortschreibung auf Grundlage des Zensus 2011.* Data available at https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/AlteAusgaben/Bevoelkerungs-fortschreibungAlt.html, accessed on August 14, 2017.

Destatis. 2015c. *Bevölkerung und Erwerbstätigkeit 2013: Vorläufige Ergebnisse der Bevölkerungsfortschreibung auf Grundlage des Zensus 2011.* Data available at https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/AlteAusgaben/BevoelkerungsfortschreibungAlt.html, accessed on August 14, 2017.

Destatis. 2015d. *Bevölkerung und Erwerbstätigkeit 2014: Vorläufige Ergebnisse der Bevölkerungsfortschreibung auf Grundlage des Zensus 2011.* Data available at https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/AlteAusgaben/BevoelkerungsfortschreibungAlt.html, accessed on August 14, 2017.

Destatis. 2016a. *Wanderungen zwischen Deutschland und dem Ausland 2014 nach Einzelaltersjahren und Geschlecht.* Data provided on March 13, 2016.

Destatis. 2016b. *Gestorbene 1952-2003 nach Alters- und Geburtsjahren sowie Familienstand.* Data provided on March 18, 2016.

Destatis. 2016c. *Bevölkerung am 31.12.1952-2011 nach Alters- und Geburtsjahren.* Data provided on March 17, 2016.

Destatis. 2016d. *Bevölkerung und Erwerbstätigkeit 1991-2011: Rückgerechnete und fortgeschriebene Bevölkerung auf Grundlage des Zensus 2011.* Data available at https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/Bevoelkerungsstand/Rueckgerechnete-Bevoelkerung.html, accessed on August 15, 2017.

Destatis. 2017a*. Wanderungen zwischen Deutschland und dem Ausland 2015 nach Einzelaltersjahren und Geschlecht.* Data provided on April 23, 2017.

Destatis. 2017b. *Wanderungen über die Grenzen Deutschlands 1990-2015 nach ausgewählten Staatsangehörigkeiten/ der Staatsangehörigkeit und Altersgruppen.* Data provided on April 18, 2017.

Destatis. 2017c. *Gestorbene 2004-2015 nach Alters- und Geburtsjahren.* Data provided on August 15, 2017.

Destatis. 2017d. *Bevölkerung am 31.12.2015 nach Alters- und Geburtsjahren.* Data provided on December 06, 2017.

Destatis. 2018a. *Wanderungen zwischen Deutschland und dem Ausland 2016 nach Einzelaltersjahren und Geschlecht.* Data provided on March 13, 2018.

Destatis. 2018b. *Wanderungen über die Grenzen Deutschlands 2016 nach der Staatsangehörigkeit und Altersgruppen.* Data provided on April 25, 2018.

Destatis. 2018c. *Gestorbene 2016 nach Alters- und Geburtsjahren.* Data provided on March 29, 2018.

Destatis. 2018d. *Bevölkerung am 31.12.2016 nach Alters- und Geburtsjahren*. Data provided on January 22, 2018.

Destatis. 2018e. *Lebendgeborene nach dem Alter der Mutter 1968-1969: DDR*. Data provided on February 08, 2018.

Destatis. 2018f. *Lebendgeborene nach dem Alter der Mutter 1970-1989: DDR.* Data provided on February 09, 2018.

Dickey, David A. and Wayne A. Fuller. 1979. “Distribution of the Estimators for Autoregressive Time Series With a Unit Root.” *Journal of the American Statistical Association* 74 (366): 427-31.

Dudel, Christian. 2014. *Vorausberechnung von Verwandtschaft: Wie sich die gemeinsame Lebenszeit von Kindern, Eltern und Großeltern zukünftig entwickelt.* Opladen, Berlin, Toronto: Verlag Barbara Budrich.

Eisenmenger, Matthias, Olga Pötzsch, and Bettina Sommer. 2006. *11. Koordinierte Bevölkerungsvorausberechnung: Annahmen und Ergebnisse.* Wiesbaden: Destatis.

European Union. 2015. *The 2015 Aging Report: Economic and budgetary projections for the 28 EU Member States (2013-2060). European Economy 3/2015.* Luxembourg: Publications Office of the European Union.

Fuchs, Johann and Katrin Dörfler. 2005. *Projektion des Arbeitsangebots bis 2050 – Demographische Effekte sind nicht mehr zu bremsen.* Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.

Fuchs, Johann, Doris Söhnlein, Brigitte Weber, and Enzo Weber. 2018. “Stochastic Forecasting Labor Supply and Population: An Integrated Model.” *Population Research and Policy Review* 37 (1): 33-58.

GENESIS-Online Datenbank. 2018a. *Lebendgeborene: Deutschland, Jahre, Geschlecht.* Data available at https://www-genesis.destatis.de/genesis/online/, accessed on May 04, 2018.

GENESIS-Online Datenbank. 2018b. *Gestorbene: Deutschland, Jahre, Geschlecht.* Data available at https://www-genesis.destatis.de/genesis/online/, accessed on January 10, 2018.

GENESIS-Online Datenbank. 2018c. *Durchschnittliches Alter der Mutter bei Geburt ihrer lebend geborenen Kinder: Deutschland, Jahre, Familienstand der Eltern.* Data available at https://www-genesis.destatis.de/genesis/online/, accessed on March 07, 2018.

GENESIS-Online Datenbank. 2018d. *Lebendgeborene: Deutschland, Jahre, Alter der Mutter, Geschlecht der Lebendgeborenen, Familienstand der Eltern.* Data available at https://www-genesis.destatis.de/genesis/online/, accessed on March 29, 2018.

GENESIS-Online Datenbank. 2018e. *Wanderungen zwischen Deutschland und dem Ausland: Deutschland, Jahre, Kontinente.* Data available at https://www-genesis.destatis.de/genesis/online/, accessed on May 11, 2018.

GENESIS-Online Datenbank. 2018f. *Bevölkerung: Deutschland, Stichtag.* Data available at https://www-genesis.destatis.de/genesis/online/, accessed on September 24, 2018.

Härdle, Wolfgang and Alena Myšičková. 2009. "Stochastic Population Forecast for Germany and its Consequence for the German Pension System." *SFB 649 Discussion Paper* 2009-009. Berlin: Humboldt-Universität zu Berlin.

Heidelberg Institute for International Conflict Research. 2017. *Conflict Barometer 2016.* Heidelberg: Universität Heidelberg, Heidelberg Institute for International Conflict.

Hyndman, Rob J., Heather Booth, and Farah Yasmeen. 2013. "Coherent Mortality Forecasting: The Product-Ratio Method With Functional Time Series Models." *Demography* 50 (1): 261-83.

Hyndman, Rob J. and M. Shahid Ullah. 2007. "Robust forecasting of mortality and fertility rates: A functional data approach." *Computational Statistics & Data Analysis* 51 (10): 4942-56.

International Organization for Migration. 2014. *Afghanistan: Migration Profile.* Kabul: International Organization for Migration.

Jaffe, Greg. 2009. "Link Between Iraq Violence and Troop Withdrawals Is Considered." *Washington Post,* July 05, 2009. http://www.washingtonpost.com/wp-dyn/content/article/2009/07/04/AR2009070402116.html.

James, William H. 2000. "Secular movements in sex ratios of adults and of births in populations during the past half-century." *Human Reproduction* 15 (5): 1178-83.

Keilman, Nico and Dinh Quang Pham. 2000. "Predictive Intervals for Age-Specific Fertility." *European Journal of Population* 16 (1): 41-66.

Keilman, Nico, Dinh Quang Pham, and Arve Hetland. 2002. "Why population forecasts should be probabilistic – illustrated by the case of Norway." *Demographic Research* 6 (15): 409-54.

Le Bras, Hervé and Georges Tapinos. 1979. “Perspectives à long terme de la population française et leurs implications économiques.” *Population* 34 (1): 1391-452.

Ledermann, Sully and Jean Breas. 1959. “Les dimensions de la mortalité.” *Population* 14 (4): 637-82.

Lee, Ronald D. 1993. “Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level.” *International Journal of Forecasting* 9 (2): 187-202.

Lee, Ronald D. 1998. “Probabilistic Approaches to Population Forecasting.” *Population and Development Review* 24 (Supplement: Frontiers of Population Forecasting): 156-90.

Lee, Ronald D. and Lawrence R. Carter. 1992. “Modeling and Forecasting U.S. Mortality.” *Journal of the American Statistical Association* 87 (419): 659-71.

Lee, Ronald D. and Shripad Tuljapurkar. 1994. “Stochastic Forecasts for the United States: Beyond High, Medium, and Low.” *Journal of the American Statistical Association* 89 (428): 1175-89.

Lipps, Oliver and Frank Betz. 2005. “Stochastische Bevölkerungsprojektion für West- und Ostdeutschland.“ *Zeitschrift für Bevölkerungswissenschaft (Comparative Population Studies)* 30 (1): 3-42.

Lutz, Wolfgang and Sergei Scherbov. 1998. “Probabilistische Bevölkerungsprognosen für Deutschland.“ *Zeitschrift für Bevölkerungswissenschaft (Comparative Population Studies)* 23 (2): 83-109.

Luy, Marc and Paola Di Giulio. 2006. “The impact of health behaviors and life quality on gender differences in mortality.” *MPIDR Working Paper* WP 2006-035. Rostock: Max Planck Institute for Demographic Research.

Mathews, Fiona, Paul J. Johnson, and Andrew Neil. 2008. “You are what your mother eats: evidence for maternal preconception diet influencing foetal sex in humans.” *Proceedings of the Royal Society B: Biological Sciences* 275 (1643): 1661-8.

McKenzie, Sheena. 2018. “How seven years of war turned Syria’s cities into ‘hell on Earth’.” *CNN*, March 15, 2018. https://edition.cnn.com/2018/03/15/middleeast/syria-then-now-satellite-intl/index.html.

Milewski, Nadja. 2010. “Immigrant fertility in West Germany: Is there a socialization effect in transitions to second and third births?” *European Journal of Population* 26 (3): 297-323.

Naderi, Robert. 2015. “Kinderzahl und Migrationshintergrund. Ein Vergleich zwischen Frauen türkischer Herkunft mit oder ohne eigene Wanderungserfahrung sowie Frauen ohne Migrationshintergrund in Westdeutschland.“ *Zeitschrift für Familienforschung (Journal of Family Research)* 27 (3): 322-42.

OECD. 2017. *OECD-Wirtschaftsausblick 2017/2.* Paris: OECD Publishing.

Pampel, Fred. 2005. “Forecasting sex differences in mortality in high income nations: the contribution of smoking.” *Demographic Research* 13 (18): 455-84.

Pearson, Karl. 1901. “LIII. On lines and planes of closest fit to systems of points in space.” *Philosophical Magazine* Series 6 (2:11): 559-72.

Pleitgen, Frederik. 2017. “Inside Aleppo: Residents return to rebuild Syria’s shattered city.” *CNN*, August 15, 2017. https://edition.cnn.com/2017/08/15/middleeast/aleppo-syria-rebuilds/index.html.

Pötzsch, Olga. 2016. “(Un-)Sicherheiten der Bevölkerungsvorausberechnungen: Rückblick auf die koordinierten Bevölkerungsvorausberechnungen für Deutschland zwischen 1998 und 2015.“ *WISTA – Wirtschaft und Statistik* 4/2016: 36-54.

Pötzsch, Olga and Felix Rößger. 2015. „Bevölkerung Deutschlands bis 2060: 13. Koordinierte Bevölkerungsvorausberechnung.“ Wiesbaden: Destatis.

Raftery, Adrian E., Jennifer L. Chunn, Patrick Gerland, and Hana Ševčíková. 2013. “Bayesian Probabilistic Projections of Life Expectancy for All Countries.” *Demography* 50 (3): 777-801.

Raftery, Adrian E., Leontine Alkema, and Patrick Gerland. 2014a. “Bayesian Population Projections for the United Nations.” *Statistical Science* 29 (1): 58-68.

Raftery, Adrian E., Nevena Lalić, and Patrick Gerland. 2014b. “Joint probabilistic projection of female and male life expectancy.” *Demographic Research* 30 (27): 795-822.

Schnabel, Sabine, Kristin von Kistowski, and James W. Vaupel, 2005. “Immer neue Rekorde und kein Ende in Sicht: der Blick in die Zukunft lässt Deutschland grauer aussehen als viele erwarten.“ *Demografische Forschung Aus Erster Hand* 2 (2): 3.

Swiaczny, Frank. 2016. "Demografischer Wandel und Migration in Deutschland: Diversität und Heterogenisierung der Bevölkerung." In *Handbuch Diversity Kompetenz. Band 1: Perspektiven und Anwendungsfelder*, edited by Genkova, Petia and Tobias Ringeisen, 155–72. Wiesbaden: Springer Fachmedien.

Trovato, Frank and N.M. Lalu. 1996. "Narrowing Sex Differentials in Life Expectancy in the Industrialized World: Early 1970's and Early 1990's." *Biodemography and Social Biology* 43 (1): 20-37.

United Nations. 2015. "World Population Prospects: The 2015 Revision: Methodology of the United Nations Population Estimates and Projections." *Working Paper* No. ESA/P/WP.242. New York: United Nations, Department of Economic and Social Affairs, Population Division.

United Nations. 2017. *Probabilistic projection of total population (both sexes combined) by region, subregion, country or area, 2015-2100 (thousands).* Available at https://esa.un.org/unpd/wpp/Download/Probabilistic/Population/, accessed on September 11, 2017.

Vanella, Patrizio. 2017a. "A Principal Component Model for Forecasting Age- and Sex-Specific Survival Probabilities in Western Europe." *Zeitschrift für die gesamte Versicherungswissenschaft (German Journal of Risk and Insurance)* 106 (5): 539–54.

Vanella, Patrizio. 2017b. "Age- and Sex-Specific Fertility in Germany until the Year 2040 – The Impact of International Migration." *Hannover Economic Papers* 606. Hannover: Gottfried Wilhelm Leibniz Universität Hannover, School of Economics and Management.

Vanella, Patrizio. 2018. "Stochastic Forecasting of Demographic Components Based on Principal Component Analyses." *Athens Journal of Sciences* 5 (3): 223–45.

Vanella, Patrizio and Philipp Deschermeier. 2018a. "A stochastic Forecasting Model of international Migration in Germany." In *Familie – Bildung – Migration: Familienforschung im Spannungsfeld zwischen Wissenschaft, Politik und Praxis. Tagungsband zum 5. Europäischen Fachkongress Familienforschung,* edited by Olaf Kapella, Norbert F. Schneider, and Harald Rost, 261—80. Opladen, Berlin, Toronto: Verlag Barbara Budrich.

Vanella, Patrizio and Philipp Deschermeier. 2018b. "A Principal Component Simulation of Age-Specific Fertility – Impacts of Family Policy on Reproductive Behavior in Germany." *Hannover*

Economic Papers 630. Hannover: Gottfried Wilhelm Leibniz Universität Hannover, School of Economics and Management.

Waldron, Ingrid. 1993. "Recent trends in sex mortality ratios for adults in developed countries." *Social Science & Medicine* 36 (4): 451-62.

Whelpton, Pascal K. 1928. "Population of the United States, 1925 to 1975." *The American Journal of Sociology* 34 (2): 253-70.

Zeit Online. 2015. „Deutschland setzt Dublin-Verfahren für Syrer aus." *Zeit*, August 25, 2015. https://www.zeit.de/politik/ausland/2015-08/fluechtlinge-dublin-eu-asyl.