Article

Prevalence and Economic Costs of Absenteeism in an Aging Population – A Quasi-Stochastic Projection for Germany

Patrizio Vanella [1,2,*], Christina B. Wilke [3], and Doris Söhnlein [4]

[1] Epidemiology Department, Helmholtz Centre for Infection Research, 38124 Brunswick, Germany; patrizio.vanella@helmholtz-hzi.de

[2] Chair of Empirical Methods in Social Science and Demography, University of Rostock, 18057 Rostock, Germany

[3] Chair of Economics, FOM University of Applied Sciences, 28359 Bremen, Germany; christina.wilke@fom.de

[4] Forecasts and Macroeconomic Analyses Department, Institute for Employment Research (IAB), 90478 Nürnberg, Germany; doris.soehnlein@iab.de

[*] Correspondence: patrizio.vanella@helmholtz-hzi.de

Abstract: Demographic change is leading to the aging of German society. As long as the baby boom cohorts are still of working age, the working population will also age - and decline as soon as this baby boom generation gradually reaches retirement age. At the same time, there has been a trend towards increasing absenteeism (times of inability to work) in companies since the zero years, with the number of days of absence increasing with age.

We present a novel stochastic forecast approach that combines population forecasting with forecasts of labor force participation trends, considering epidemiological aspects. For this, we combine a stochastic Monte Carlo-based cohort-component forecast of the population with projections of labor force participation rates and morbidity rates.

This article examines the purely demographic effect on the economic costs associated with such absenteeism due to the inability to work. Under expected future employment patterns and constant morbidity patterns, absenteeism is expected by close to 5 percent by 2050 relative to 2020, associated with increasing economic costs of almost 3 percent. Our results illustrate how strongly the pronounced baby boom/ baby bust phenomenon determines demographic development in Germany in the midterm.

Keywords: Cohort-Component Method; Multivariate Methods; Time Series Analysis; Monte Carlo Methods; Stochastic Forecasting; Demography; Statistical Epidemiology; Labor Market Research; Health Economics

1. Introduction

The demographic transition is leading to the aging of German society [1,2]. As long as the baby boomers in Germany are still of working age, the workforce will also age - and shrink and rejuvenate as that generation gradually reaches retirement age [3]. The 1964 cohort turned 55 in 2019 and is expected to regularly retire at age 67 in 2031 [4]. In the short term, therefore, the German economy will need to focus on the aging of the workforce while in the medium and long term the shrinking and rejuvenation process will prevail [1]. An important issue to address, among others, is how this demographic development will affect the future economic costs of companies and the German health insurance system with respect to absenteeism.

Absenteeism (periods of health-related incapacity to work) is a topic yet under-investigated when it comes to its economic costs. Although aging is a well-researched topic in the contexts of labor supply [3], pensions [4], healthcare (costs) [5], or long-term care and disabilities [6], the question of absenteeism going along with aging has yet been given

much less attention. It is, after all, not only a question of how strong the labor supply will be affected by aging but also to what extent it is threatened to be temporarily lost due to (chronic) diseases and illnesses. The present aging report by the European Union (EU), for instance, does not even mention the problem of absenteeism [5]. Studies addressing absenteeism, on the other hand, remain rather descriptive, not offering projections of absenteeism in the context of the demographic transition – although it is well-known that times of absenteeism occur more often and for a longer period later in work life [7].

In Germany, which is the most populous country in the EU [5], aging is a topic of major concern given its low fertility [8] and decreasing mortality [9] for almost half a century. At the same time, since the mid-2000s, a trend toward increasing absenteeism can be observed, as the average number of days of absenteeism per case increases with age [10]. Is this trend likely to continue in the future? What economic costs are already incurred by companies because of absenteeism, and what costs can ceteris paribus (c.p.) be expected in the future as a result of the German demographic development? These are the questions we address in this paper.

How the working-age population will develop in the future depends on the underlying population trends. These are determined by the three major demographic components fertility, migration, and mortality [1]. Of particular relevance for the development of the labor force are the fertility rates, which have a delayed effect on the labor force when the new-born have entered the working age [8], and migration, which affects the labor market relatively quickly as people tend to migrate when they are of working age. However, to what extent migrants also succeed in entering the German labor market, very much depends on the migrants' sociodemographic background [3,11]. For instance, the labor market integration of refugees takes significantly more time on average than for foreign citizens from the EU [11]. Mortality trends, on the contrary, have a rather negligible impact on the labor force, as they become quantitatively noticeable rather at older ages, long after the end of the common working life [3]. In this paper, we will project future labor force development driven by these components.

A first forecast model for absenteeism in the context of aging and the effect on the associated economic costs was suggested by Wilke [10], also for the German case. The model is based on age-specific morbidity risks derived from data provided by the Federal Institute for Occupational Safety and Health ($BAuA$) [12]. In our contribution, we further elaborate this approach by connecting it to a stochastic population forecast for Germany, developed by Vanella and Deschermeier [1]. Moreover, we take trends in labor force participation into account to acknowledge their potential impact on absenteeism. For this, we implement updated forecasts of age- and sex-specific labor force participation rates (LFPRs) in our model, which have been conducted by Fuchs and colleagues based on an own stochastic forecast model [3,13].

Section 2 outlines the data and the methodological approach used in our study in more detail. Section 3 shows the results for our projection of future labor force development and corresponding cases of absenteeism until the year 2050 – provided age-specific employment and disability rates remain constant. Subsequently, these purely demographic effects on the future development of absenteeism are evaluated further in terms of the associated economic costs. In Section 4, we then discuss limitations and potential improvements and close with an outlook on questions for further research.

2. Materials and Methods

2.1 Stochastic Population Forecast

As a baseline of our projection, we compute a stochastic population forecast by age and gender for Germany over the period 2021-2050, following a stochastic cohort-component approach for age- and sex-specific population forecasting suggested by Vanella and Deschermeier [1] with an adjustment to the migration forecast. There, the authors forecast the population based on stochastic principal component-based time series methods for the major demographic components fertility, international net migration, and mortality.

Stochasticity is considered in the model by Monte Carlo simulation (10,000 draws) of all variables, resulting in 10,000 trajectories of the future population by age and gender. Our migration model, however, differs from that suggested in the mentioned study. For Germany, migration rates are more stable than flows, as they consider the baseline population [11]. Moreover, for the case of small-area migration in Germany, Vanella et al. [14] show that a naïve random walk forecast of pseudo migration rates gives better fits to historical migration patterns in Germany. In that vein, we use the latter approach, adjusted to national age- and sex-specific pseudo net migration rates for Germany.

Hence, in the *first* step of the analysis, we forecast the future birth numbers. For this, we multiply forecasts of the age-specific fertility rates (ASFRs), as derived according to Vanella and Deschermeier [1,8], with simulations of the female population in the corresponding fertile age. I.e., in trajectory t, the births in year y $(B_{y,t})$ are the scalar product of the vectors of the simulated ASFRs and the corresponding female population

$$B_{y,t} = \sum_{a=15}^{50} \varphi_{y,a,t} * P_{y-1,a-1,f,t},$$ (1)

with $\varphi_{y,a,t}$ being the ASFR of females aged a years in trajectory t in year y and $P_{y-1,a-1,f,t}$ being the simulated female population in age $a-1$ at year-end $y-1$. Therefore, the ASFR is estimated on the female population at the end of the previous calendar year. As there are no clearly defined lower or upper limits of the reproductive age, 15 is defined as the lower limit, and age group 50 refers to the baseline population 50-54 years of age. Therefore, we assume the same ASFR for females aged 50-54. These assumptions, based on the available data, give plausible age schedules for fertility and, therefore, offer a good basis for forecasting the total births. The baseline birth data have been partly (1968-1991) provided on request by the German federal statistical office (Destatis) for the two former German states to the first author of the present paper for earlier studies [15-18], and partly (1992-2020) been downloaded for unified Germany from Destatis' database GENESIS-Online [19]. The denominators of the ASFRs (and all other rates used for the study) are the corresponding end-of-year population estimates, which are taken from the Human Mortality Database (HMD) for all years until 2017[1] [20-22]. The HMD estimates derived by Klüsener et al. [23] are, especially for the years before 2011, preferable to the original Destatis data, as they are adjusted to errors arising from the population updating in intercensal periods and errors from unobserved migration, which particularly bias the old-age population estimates [1]. Since at the time of writing this paper no data beyond 2017 was available on the HMD, we rely on original Destatis population data, provided on request to us [24], for the years 2018-2020.

Second, after deriving the birth numbers, we simulate the share of males (and indirectly females), respectively, among all births[2] by time series ARIMA methods, following Vanella and Deschermeier [1]. Gender shares among the births since 1950 are downloaded from GENESIS-Online [25]. Male births are consequently computed by multiplying the share of males in year y in trajectory t $(r_{y,t})$ with all births in the said year and trajectory:

$$B_{y,m,t} = r_{y,t} * B_{y,t}.$$ (2)

The female births are, consequently,

$$B_{y,f,t} = (1 - r_{y,t}) * B_{y,t} = B_{y,t} - B_{y,m,t}.$$ (3)

---

[1] Note that the HMD estimates refer to the January 1st rather than December 31st. Therefore, for instance, the population estimates by the HMD for 2018 correspond to Destatis estimates for 2017.

[2] The official statistics do not differentiate further genders but distributes them to either males or females. Accordingly, our forecast sticks to the binary gender definition.

Third, we simulate migration in- and outflows by age and gender, generalizing the suggestion by Vanella et al. [14]. Migration flows are not available by cohort but only completed years of age of the migrants. We, therefore, assume migration and births to occur uniformly over the year[3]. These assumptions allow us to approximate cohort-based age- and sex-specific migration flows based on the migration flows by completed age through averaging. I.e., based on the assumption that 1/12 of all annual migrations happen each month, we can conclude that half of the annual migrations have happened by the end of each year. Therefore, the average age of a migrant is exactly between two adjacent birthdays. For instance, among female migrants of age 30 in the year 2020, one-half is assumed to have been born in the calendar year 1990, while the other half consequently has been born in 1989. Based on these flow data, we compute age- and sex-specific emigration rates and pseudo-immigration rates, following Fuchs et al. [11]. Migration flows by age and gender for the years 1991-1999 had been provided to the first author of the present paper by Destatis for an earlier study [27], the corresponding data for the period 2000-2020 are available via GENESIS-Online [28]. Based on the migration flow data in connection to the earlier mentioned population estimates, we derive (pseudo-) migration rates, which we forecast according to Vanella et al. [14]. The (immigration) emigration of individuals of age a and gender g in year y and trajectory t, therefore, is given by

$$I_{y,a,g,t} = \begin{cases} i_{y,0,g,t} * B_{y,g,t} \ for \ a = 0, \\ i_{y,a,g,t} * P_{y-1,a-1,g,t} \ for \ a > 0 \end{cases} \tag{4}$$

and

$$E_{y,a,g,t} = \begin{cases} e_{y,0,g,t} * B_{y,g,t} \ for \ a = 0, \\ e_{y,a,g,t} * P_{y-1,a-1,g,t} \ for \ a > 0' \end{cases} \tag{5}$$

for immigration and emigration, respectively. Here, $i_{y,a,g,t}$ ($e_{y,a,g,t}$) is the pseudo-immigration (emigration) rate of individuals aged a of gender g in year y, and trajectory t. Note that migration estimates here are computed as the product of the pseudo-immigration (emigration) rate and the end-of-year population of the corresponding cohort in the previous year and the target (origin) country. To include migration among newborns, we refer to the births occurring in year y as the denominator. Following Vanella et al. [14], we compute pseudo-net migration rates by age and gender as the difference of pseudo-immigration rate and emigration rate of the same stratum, i.e. $n_{y,a,g,t} := i_{y,a,g,t} - e_{y,a,g,t}$. Vanella et al. [14] show that for Germany, a naïve multivariate random walk approach based on migration rates performs best to predict migration future migration flows. We adjust their approach by assuming the first differences of the pseudo-net migration rates $\Delta n_{y,a,g,t} := n_{y,a,g,t} - n_{y-1,a,g,t}$ to follow a multivariate Gaussian distribution:

$$\begin{bmatrix} \Delta n_{y,0,m,t} \\ \Delta n_{y,1,m,t} \\ \vdots \\ \Delta n_{y,0,f,t} \\ \vdots \end{bmatrix} \sim \mathcal{N}(\vec{0}, \boldsymbol{\Sigma}) \ \forall \ y, t, \tag{6}$$

with $\vec{0}$ being a 2a-dimensional (in our case 192-dimensional) null vector and $\boldsymbol{\Sigma}$ being the empirical covariance matrix of the first differences of all age- and sex-specific pseudo-net migration rates for 1995-2020. Monte Carlo sampling from (6) leads to 10,000 trajectories of the age- and sex-specific pseudo-net migration rates. Multiplying those with the population bases according to (4) and (5), we can simulate the distributions of net migration for each stratum and year as

$$N_{y,a,g,t} = \begin{cases} n_{y,0,g,t} * B_{y,g,t} \ for \ a = 0, \\ n_{y,a,g,t} * P_{y-1,a-1,g,t} \ for \ a > 0 \end{cases} \tag{7}$$

---

[3] Which in praxis is not the case. For births, for instance, there is strong annual seasonality [26].

Fourth, we compute the deaths according to Vanella and colleague [1,9]. For this, age- ad sex-specific survival rates are computed retrospectively as quotients of survivors of some cohort at the end of the year and the sum of the survivors and deaths among the same cohort at the same point in time. This way, we include mortality among the children born in the current year and migration into our mortality estimates, including the timing of births and migration as represented by the historical data. The simulation of survival rates combined with the population estimates for the previous year and the migration flows and births simulated according to (3)-(7) gives the death simulations:

$$D_{y,a,g,t} = \begin{cases} (1 - s_{y,0,g,t}) * B_{y,g,t} \text{ for } a = 0, \\ (1 - s_{y,a,g,t}) * (P_{y-1,a-1,g,t} + N_{y,a,g,t}) \text{ for } a > 0' \end{cases} \tag{8}$$

with $s_{y,a,g,t}$ being the survival rate of individuals of age a and gender g in year y and trajectory t. The reciprocal $(1 - s_{y,a,g,t})$ is the corresponding mortality rate. The data underlying the computation of the survival rates are, next to the earlier mentioned population estimates, data on annual deaths by cohort and gender, provided on several occasions on request by Destatis to the first author [29,30].

Fifth, combining the results from (3)-(8), we obtain the cohort-component simulations of the end-of-year population as

$$P_{y,a,g,t} = \begin{cases} B_{y,g,t} + N_{y,0,g,t} - D_{y,0,g,t} \ for \ a = 0, \\ P_{y-1,a-1,g,t} + N_{y,a,g,t} - D_{y,a,g,t} \ for \ a > 0. \end{cases} \tag{9}$$

2.2 Projection of Labor Force Participation in the Context of Increasing Retirement Ages

Sixth, we combine our population forecast obtained from (9) with an updated stochastic forecast of labor force participation rates (LFPRs) suggested by Fuchs et al. [3], which includes trends in age-, gender- and nationality-specific labor market participation. These forecasts are especially crucial for our research question, as trends of increasing LFPRs are observable in the long run among females [3] and we will see increasing LFPRs among the elderly as a consequence of demography-related pension reforms, which cause legal retirement ages to increase until 2031 [4]. The updated LFPRs are provided by Fuchs et al. [13]. The authors, therefore, include trends in labor force participation by migration (for instance, longer periods of labor market integration) in their forecast. Therefore, we implicitly assume shares of migrants in the population according to Fuchs et al. [13], as our model does not predict the foreign separately from the German population. The projection of labor supply by age group and gender can then be derived by

$$L_{y,a,g,t} = l_{y,a,g} * P_{y,a,g,t}, \tag{10}$$

with $l_{y,a,g}$ being the median forecast of the LFPR of individuals of age a and gender g for year y according to Fuchs et al. [13]. Note that we include the LFPRs only deterministically in the model, as they are subject to many aspects, such as the socio-economic composition of the population or the overall economic development. Since our focus is to investigate the pure demographic effect on future absenteeism, a deterministic inclusion of labor market effects fits our purpose. Including at least the expected trends in labor force participation appears necessary, however, to give a realistic estimate of the demographic effect under increasing retirement ages. We only adjust the LFPRs insofar, as Fuchs et al. [13] simulate five-year age groups. Our model is based on single years of age and we want to avoid hard cuts in the LFPR curves, which are a consequence of larger age groups. Therefore, we assume age- and sex-specific LFPRs (ASSLFPRs) as provided in Fuchs et al. [13] for age groups for the median age of each age group (e.g., for age group 25-29, the LFPR from Fuchs et al. is assumed for age 27 in our model), and interpolate the LFPRs for the ages between the knots by natural cubic splines [31]. Yet, our aim is not to give a forecast of the future labor supply but to include the expected effects of labor supply

change in our model of future absenteeism. Since labor supply does not equal the population in labor, the labor force projections are just used as auxiliary data that include the effect of delayed retirement in our projection. Assuming that the share of the labor force that is employed remains constant, the relative change between periods y and $y + \tau$ in employed persons for each age-gender stratum under changing LFPRs and the demographic trends in trajectory t derived from (9) can then be computed by

$$\ell_{y,\tau,a,g,t} := \frac{L_{y+\tau,a,g,t}}{L_{y,a,g,t}}. \tag{11}$$

2.3 Projection of Relative Increase in Absenteeism given Demographic and Economic Trends

Seventh, we combine our projections of the labor force effect according to (11) with information on absenteeism by age group to deliver stochastic estimates of the demographic effect on absenteeism. We do not know the exact number of days of absenteeism. The most informative data available publically are annual reports on work-related safety and health issues, provided by the Federal Ministry of Labour and Social Affairs (*BMAS*) and the BAuA. We specifically use the latest report for the year 2019 [7]. In Germany, every individual is health insured by either social health insurance (GKV) or private health insurance (PKV) companies. In our reports, the data are restricted to GKV insured; age-specific data for PKV insured are not publically available. We use estimates for both cases of absenteeism, standardized to 100 years of insurance membership, and averages of days per case are reported for 5-year age groups. The numbers are given in Table 1.

Table 1. Standardized age-specific Cases of Work Absence per 1 Year of Full Insurance and Average Length of Absence as Days for GKV members in 2019 (Sources: [7]; authors' computation and illustration)

Age Group	Annual Cases per Capita	Average Days per Case	Average Annual Days per Capita
15-19	2.57	5	12.85
20-24	2.1	6	12.6
25-29	1.65	8	13.2
30-34	1.6	9	14.4
35-39	1.6	10	16
40-44	1.53	11	16.83
45-49	1.48	13	19.24
50-54	1.53	15	22.95
55-59	1.65	17	28.05
60-64	1.74	21	36.54
65+	0.71	23	16.33

The number of cases does not increase monotonically by age, as knowledge about the connection between age and morbidities would suggest (see, for instance, [6] on disability risks in the context of long-term care insurance). Instead, we have more of a bathtub shape in the age groups below 65, followed by a sharp decrease for the elderly. The average sick days per case, however, show an increasing duration of absenteeism with increasing age. Multiplying these two statistics gives an average of days of absenteeism per year, standardized to one year of full insurance as an employed person in the GKV, which is given in the last column of Table 1. There, we see clear trends of longer absenteeism for age over 24. For the teenage group, we see a slightly higher average absenteeism in comparison to the 20- to 24-year-olds. This might be a rounding error resulting from the rounding in the available data (note that the average of days per case is rounded to integers). An alternative explanation might lie in more absenteeism as a result of more injuries due to riskier

behavior, e.g., observable in higher traffic accident rates [32]. The dip in the elderly is un-intuitive but could be explained by a positive selection of healthier workers employed in less exhaustive fields and having a higher work motivation, therefore, in many cases working beyond their individual retirement age, following Wilke [10].

It has to be noted that the data are not just a sample of all insured workers but somewhat biased. Age structure [33] and morbidities of the insured diverge, with the PKV attracting, on average, wealthier and healthier customers [34], since insurance companies offer more individualized contracts [35] and, in some regions, a more extensive healthcare supply is available for PKV in comparison to GKV insured [36]. Therefore, our data will probably overestimate the morbidities of the overall population. To circumvent this limitation in the data, we do not directly address the days of absenteeism but simply assume the rela-tion of sick days between the age groups, i.e. the risk of sickness, to remain constant over time. Moreover, we assume constant shares of members in the GKV and PKV for the pro-jection. Under these assumptions, we can project the relative change in absenteeism over time, restricted to future developments in the population size and structure and labor force participation, including increases in incidences in the oldest age group caused by increases in legal retirement ages which will occur until 2031 [4]. Yet again, we take the age group-specific numbers from Table 1 and interpolate them by natural cubic splines, as performed for the LFPRs. Our sickness data does not discriminate by gender; therefore, we will assume the morbidities to depend exclusively on age, not gender. Indeed, differ-ences in morbidities between the genders exist, yet appear more subtle in the labor age groups, becoming more crucial among the elderly [6,37]. Therefore, the error resulting from this assumption appears negligible. In particular, we take (11), which includes de-mographic and economic trends and multiply with average annual per capita absenteeism $(d_{y,a})$, which then serve as some kind of weighting factor of age-related morbidity:

$$S_{y,a,g,t} := d_{y,a} * L_{y,a,g,t}, \tag{12}$$

Where $S_{y,a,g,t}$ is the annual sick days of individuals of age a and gender g in trajectory t and in year y, which includes both demographic effects (stochastic) and labor market effects (deterministic), applying a status quo assumption to epidemiological trends. As $d_{y,a}$ is biased [10], we opt to derive the relative change in absenteeism instead of absolute numbers. Assuming that the relative bias of age-specific absenteeism is the same for all age groups, the biases will cancel out in a relative measure. The sum of absenteeism in year y over all demographic groups is $S_{y,t} = \sum_a \sum_g S_{y,a,g,t}$. The absenteeism in year y, relative to the year 2020, in each trajectory is then

$$\kappa_{2020,y,t} := \frac{S_{y,t}}{S_{2020}}. \tag{13}$$

2.4 Projection of Relative Increase in Economic Costs by Absenteeism Trends

In the *eighth* and final step, we project how, based on our previously derived results, ab-senteeism will affect health economics. Given the derived demographic development and combined with the assumed economic and epidemiological trends, we project costs caused by productivity loss, ceteris paribus. For this, we borrow Wilke's [10] estimates of age-specific productivity, measured in € of 2018, and loss thereof for each day of absen-teeism. The most important numbers are given in Table 2.

Table 2. Annual gross income and productivity loss per day of absenteeism 2018 (Source: [10]; own illustration)

Age Group	Average Gross Income [as €]	Loss of Productivity by Day [as €]
20-24	21,246	58.21
25-29	31,790	87.10
30-34	39,826	109.11
35-39	43,083	118.04
40-44	45,610	124.96
45-49	46,075	126.23
50-54	45,972	125.95
55-59	43,689	119.70
60-64	40,853	111.93
65-69	16,233	44.47

Again, we take the last column and interpolate the values as described earlier. We can then estimate the relative change in economic costs of absenteeism by multiplying (12) with the average daily loss of productivity for the corresponding demographic stratum, say $c_{y,a}$:

$$C_{y,a,g,t} := c_{y,a} * S_{y,a,g,t}.\qquad(14)$$

Similar to (13), we compute the change in economic costs by absenteeism, relative to 2020 by

$$\pi_{2020,y,t} := \frac{C_{y,t}}{C_{2020}},\qquad(15)$$

with $C_{y,t}$ being the total costs of absenteeism calculated over all age groups in year y and trajectory t.

We will present the results of our simulations in the next section.

3. Results

Figures 1 and 2 illustrate the results of our population forecast. We show the forecasts until 2050, aggregated to ten-year age groups. As our focus is on the labor age population, we limit our analysis to age groups 15-74.[4] The black, continuous lines show the past observations since 1990, the median forecast is visualized as blue dashed lines and the 75% prediction intervals (PIs) are added as violet, dotted lines. Those have been derived from Monte Carlo simulation via cohort-component forecasting as described in 2.1. The interested readers find forecast results for the demographic components in Appendix A, and more detailed results generated by the simulations are added in Online Supplement A.

---

[4] The model simulates ages by year 0-99, and 100+ as a cumulative age group.
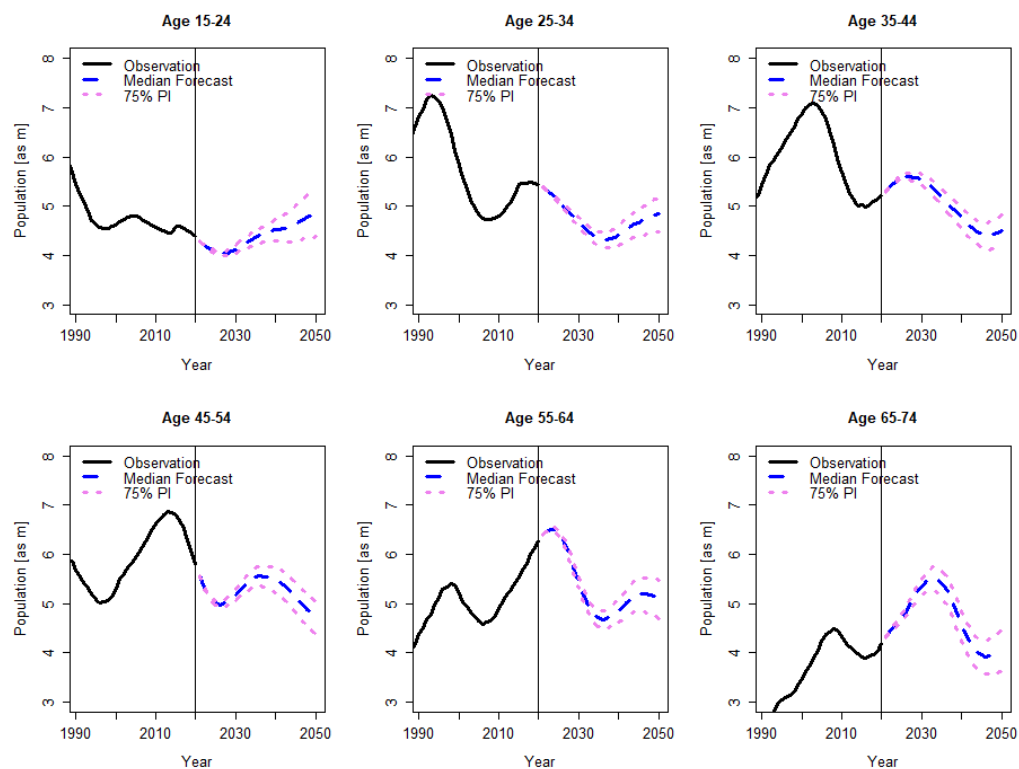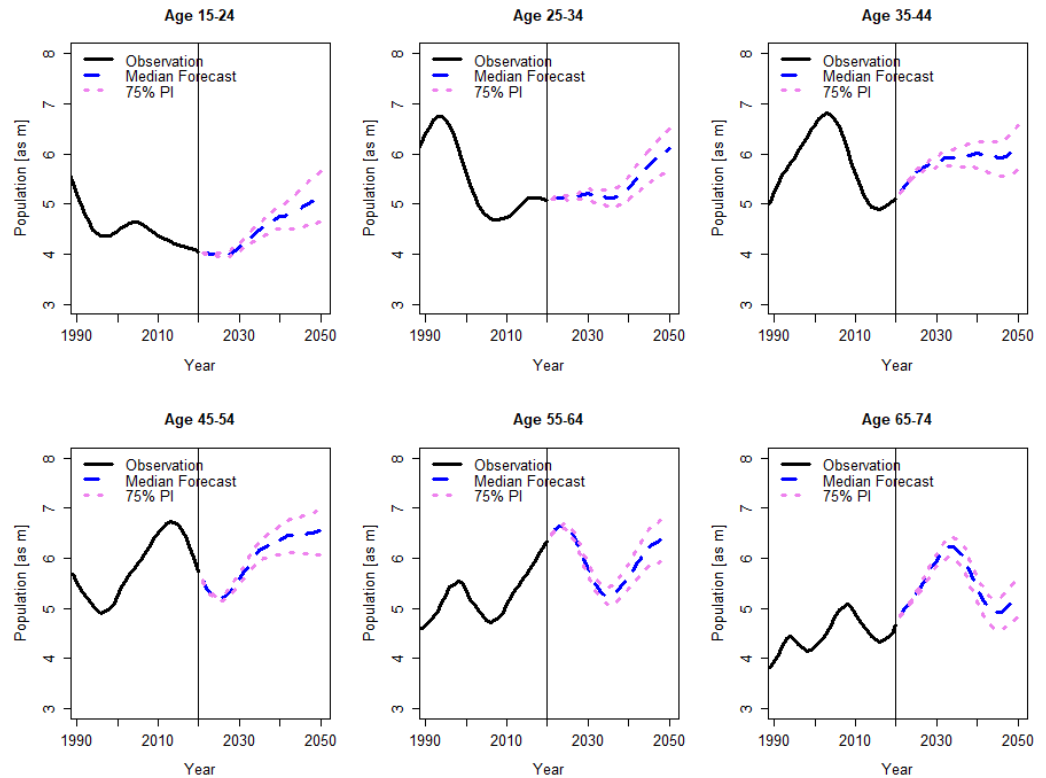
Figure 1. Forecasts of Male Population in Labor Age by 10-Year Age Groups with 75% Prediction Intervals (Sources: [21]; authors' computation and illustration)

We expect a dip in the younger population by the end of the current decade, which is a consequence of a long period of very low fertility rates in Germany (the total fertility rate in Germany has consistently been below 1.5 children between 1981 and 2014 [8]). As a consequence of increasing fertility rates and positive net migration in the young age groups, this trend will invert in the long term, however. Volatility in the youngest age groups increases significantly at the end of the 2030s, which is because a large share of that age group is not yet born. An apparent trend is a large wave we see in the age group 55-64 in the 2020s and, correspondingly, for the age group 65-74 in the following decade, which is the strong baby boom cohort, whose labor force participation is of special interest for our research question. The wave is more distinct for the females than for the males, which is due to the lower mortality of females, which becomes more emphatic in those age groups [9].

Figure 2. Forecasts of Female Population in Labor Age by 10-Year Age Groups with 75% Prediction Intervals (Sources: [15]; authors' computation and illustration)

The latter point stresses the importance of labor force participation trends by females. LFPRs of females traditionally are, on average, smaller than for males, as the females have been taking over more responsibilities in the household, and childcare opportunities have not been sufficient to allow for both parents to participate in the labor market. Since the late 1970s, more females have had the opportunity and will to follow a career of their own [8]. As a consequence, LFPRs have been increasing, which is a trend expected to persist in the future [3]. Figure 3 shows the LFPRs by age group and gender, nowcast by Fuchs et al. [13], for 2020 against their median trajectory for 2050. Our cubic splines interpolations are included in the figure as lines. The authors forecast increases in the LFPRs, especially among females. Under increasing trends in tertiarization and corresponding longer periods of education [38] that are connected to later entries in the labor market.

Figure 3. Age- and Sex-specific Labor Force Participation Rates for 2020 and 2050 (Sources: [25]; authors' computation and illustration)

Figure 4 shows our projection of the percentage change in overall absenteeism in Germany in relation to 2020, derived from our model, with 75% PIs. It has to be stressed that stochasticity in the model draws from demography only. Labor force participation is assumed as illustrated in Figure 3, and prevalence rates are assumed constant as derived by Wilke [10] and our interpolation of her estimates. Therefore, we exclusively estimate risk arising from the demographic developments, assuming fixed trends in labor force participation and epidemiology. Interestingly, we can expect waves in absenteeism with increases over the next years, as the baby boomers will be in the late phase of their employment period, which is associated with relatively high morbidity and connected phases of absenteeism (see Table 1). This development is enhanced by the mentioned increases in legal retirement ages until 2031. The delaying retirement effect of those increases will then stabilize, which will lead to further increases in old-age pensions [4], whereas decreases in absenteeism can be expected. During the 2040s, increases in absenteeism can be expected as a consequence of the high net migration Germany has witnessed since the 2010s, especially among young migrants (see, for instance, Figures A3 and A4). Those having immigrated at a young age in the strong migration waves will in the 2040s reach their late labor age, which will be associated with high prevalences of absenteeism.
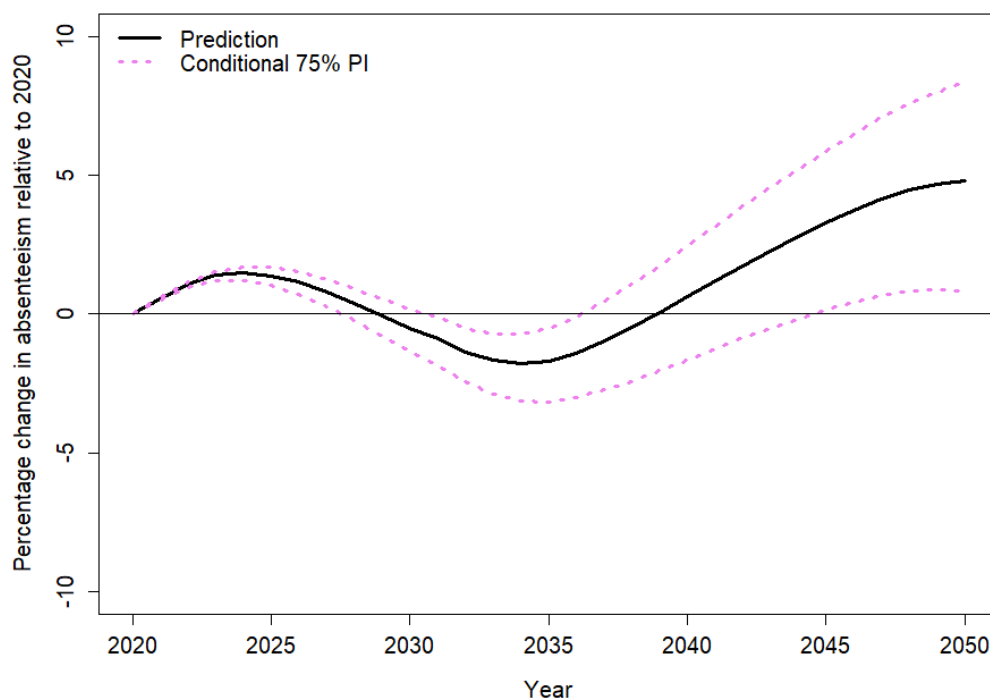
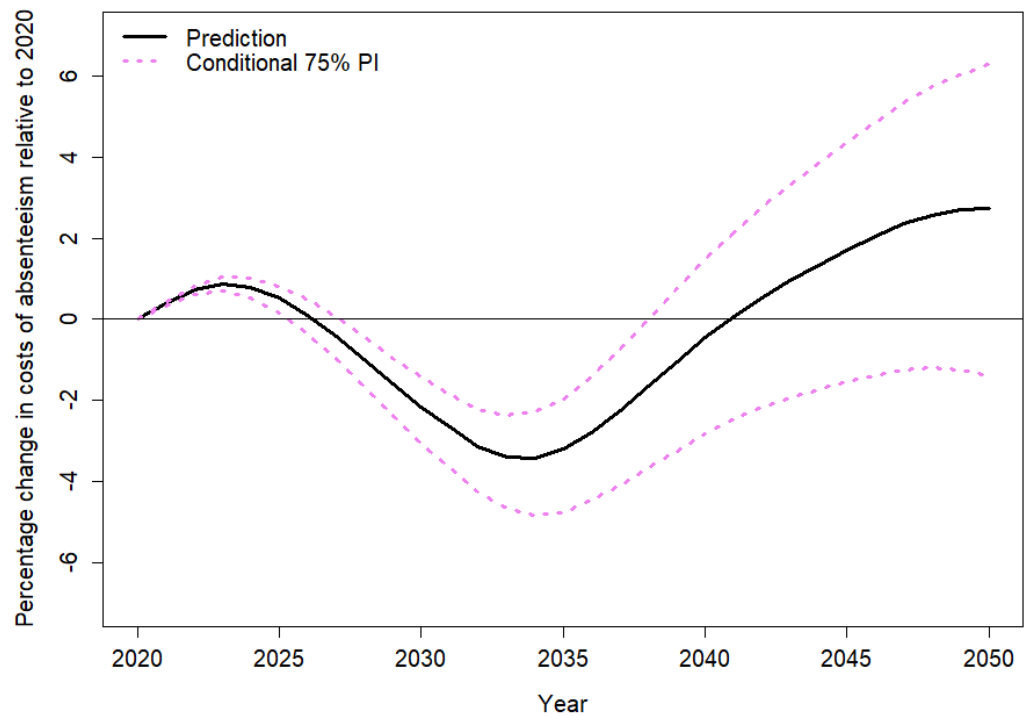**Figure 4.** Projection of relative percentage change in overall annual absenteeism to reference year 2020 (Sources: [7,13]; own computation and illustration)

Finally, we consider the relative change in economic costs due to absenteeism associated with the demographic development, illustrated in Figure 5. Whereas the overall trending of the curve looks similar to Figure 4, we see less distinct waves, as not only the overall days of absenteeism factor into this but also the loss of productivity, measured according to Table 2. This, therefore, appears to give a more realistic picture of the economic consequences of the aging process, since we account for the worth of an employee to their employer as well, not just if they are at work or not.

Figure 5: Projection of relative percentage change in overall annual costs of absenteeism to reference year 2020 (Sources: [7,10,13]; own computation and illustration)

4. Discussion

We suggest a forecast model of absenteeism, using Germany as a case study and projecting absenteeism both in days and costs until 2050. We show that absenteeism is a complex and, for aging economies, severe topic. Yet, it is rather under-investigated in the scientific literature. Absenteeism depends not only on demographic development but also on socio-economic trends, such as labor force participation, as well as epidemiological trends, such as trends in morbidity prevalence and working environments.

We focus on the first part in most detail, employing an adjusted version of an established stochastic population forecast model for Germany, suggested by Vanella and Deschermeier [1], including labor force participation rates and morbidity rates deterministically. LFPRs are assumed according to a recent forecast by Fuchs et al. [13], whereas age-specific morbidity rates are held constant as under the status quo scenario. We believe that our model thus offers a new approach towards research in economic and demographic forecasting with great potential for further development and improvement.

Still, our model has some limitations. First, the model relies on a forecast of the labor supply potentially available to the labor market, based on a stochastic population forecast in combination with a forecast of labor force participation rates. This, however, does not necessarily correspond to the number of persons being employed as we do not model the demand side of the labor market. To do this, another sophisticated forecast of the future economic situation would be required, which is not only beyond the scope of this paper as

such but also difficult to project in the long run. Moreover, additional data on the education and training of the working-age population would be necessary and needed to be matched with projected labor demand.

Apart from this, a higher differentiation of absenteeism, e.g. by gender or nationality, would improve our projections further. However, this type of data is not publicly available.

In addition, our projections of days of absenteeism and associated economic costs are not to be mistaken as true forecasts (for a distinction between forecasts and projections, see, for instance, [39]), as we provide forecasts on the population but include labor force participation rates deterministically, only.

Moreover, we do not conduct a forecast of the relevant epidemiological parameters, i.e. we do not make any assumptions about how age-specific absenteeism might develop in the future. In an aging population, so far, empirically, it remains unclear, whether overall morbidity will decrease or increase. In the literature, two opposing theories can be found: the theory of morbidity compression and the medicalization theory. While the first theory assumes morbidities to be delayed to later points in time, quasi-parallel to increases in life expectancy [40], the latter assumes increases in life expectancy to be completely spent in poor health [41]. While there is some evidence that the truth lies somewhat in between these two extreme scenarios [6], little is yet known about age-specific morbidity patterns. We tried forecasting the epidemiological trends ourselves in gathering time series data for age-specific absence rates from all previous BMAS/BAuA reports, however, we found significant structural breaks in 2016, not allowing a sound basis for future projections. Therefore, we decided to follow a simple but well-defined projection approach where epidemiological risks are assumed to remain constant. Changes in absence risks in the model are thus exclusively rooted in changes in age-specific LFPRs and therefore only depicture the pure demographic effects.

Regarding days of absenteeism and associated economic costs, our data was restricted to GKV data, which might be strongly biased, as has been discussed earlier. To account for this, we do not deliver absolute results for days of absenteeism and associated economic costs but only relative changes over time, so that biases level off.

Last, our model only relies on pre-pandemic data. Therefore, our forecast does not include any influence the COVID-19 pandemic might have on demographic or epidemiological trends. For our model, however, it is only the working-age population that is relevant. Vanella et al. [42,43] have shown in international studies on excess mortality and case fatality risks that mortality trends in these age groups so far have not been significantly tampered by COVID-19. Therefore, the underlying mortality forecast is robust. For fertility in Germany, Vanella et al. [26] did not find statistical evidence for any effect of the pandemic. Therefore, our birth forecast holds as well. International migration flows slumped during the pandemic. So far, however, the effect of the pandemic on international migration is unpredictable as reliable data is still missing. The pandemic in 2020 halted the rise in employment rates that has been continuing since 2006. Even though LFPRs collapsed in the first months of the pandemic, a rising trend is anticipated again once the pandemic situation has ended. A similar trend is expected for 2021 [44]. Germany introduced the first major contact reduction measures in March 2020 [45]. Overall, the short-term effects of the pandemic can be explained quite well. For example, well over half of the significant increase in transitions from employment to unemployment in April 2020 was probably due to the shutdown. The effects were even more pronounced in the absence of new hiring [46]. It is not yet clear, however, whether and to what extent long-term negative effects of the pandemic are to be expected. With or without the pandemic, in the medium and long run, a decline in the working-age population will be inevitable given the German demographic development [3].

Keeping these limitations in mind, our model provides a solid first step towards under-standing absenteeism and its economic costs given the ongoing aging process. Further-more, we show that increases in the statutory retirement age do not lead to one-to-one increases in LFPRs but are also, as shown by Vanella et al. [4], associated with increases in disability pensions and higher absenteeism. Common projections on demography and pensions typically do not take these aspects into account (see, for instance, Vanella et al. [4], for an overview). Our approach, therefore, could be included in projections of future labor markets and pensions as a baseline for economic planning and political decision-making.
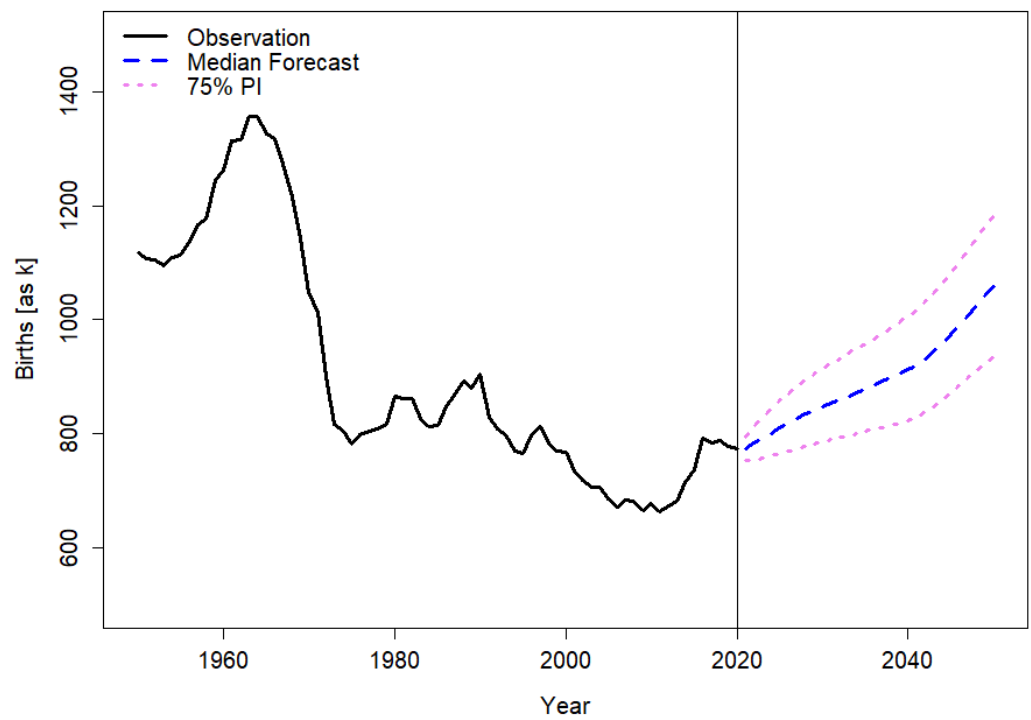
Appendix A. Supplementary Figures

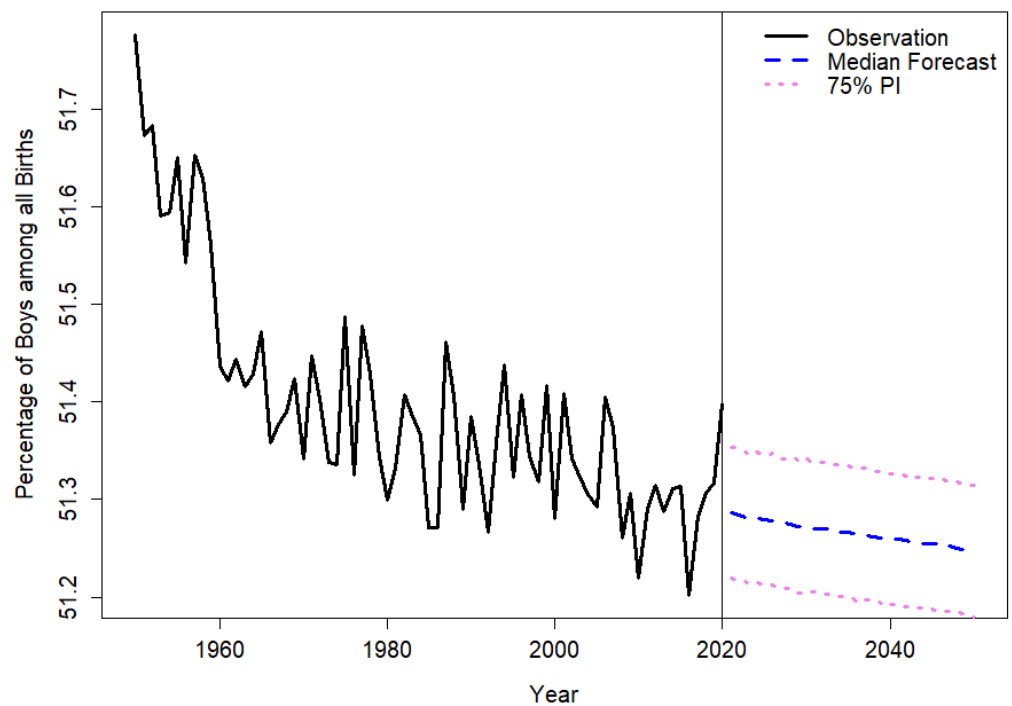**Figure A1.** Births Forecast with 75% Prediction Interval (Sources: [19]; authors' computation and illustration)

Figure A2. Forecast of Males among all Births with 75% Prediction Interval (Sources: [19]; authors' computation and illustration)
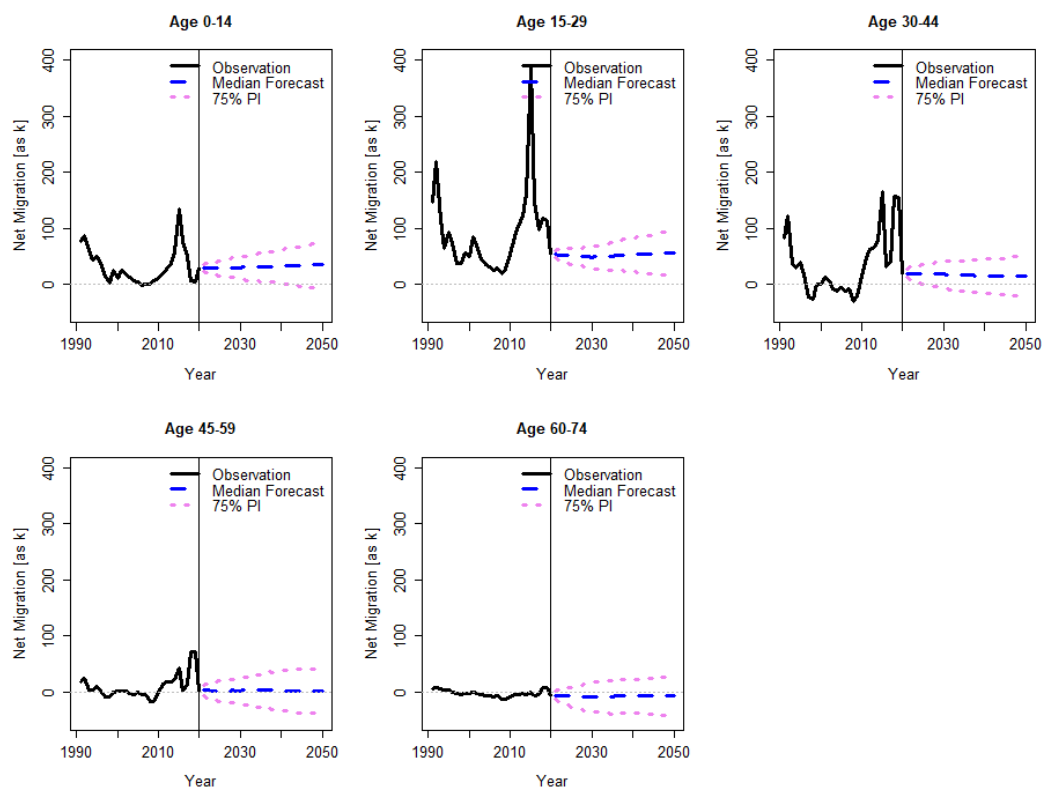
Figure A3. Forecasts of Net Migration of Males by 15-Year Age Groups with 75% Prediction Intervals (Sources: [27,28]; authors' computation and illustration)
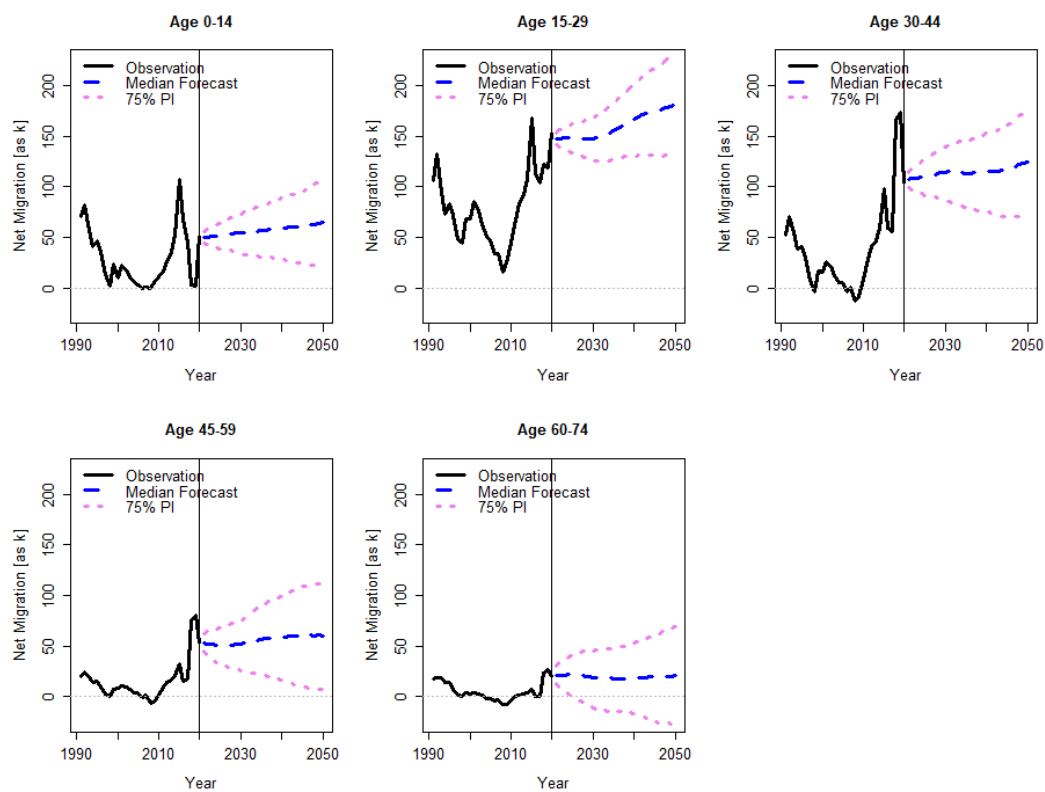

Figure A4. Forecasts of Net Migration of Females by 15-Year Age Groups with 75% Prediction Intervals (Sources: [27,28]; authors' computation and illustration)
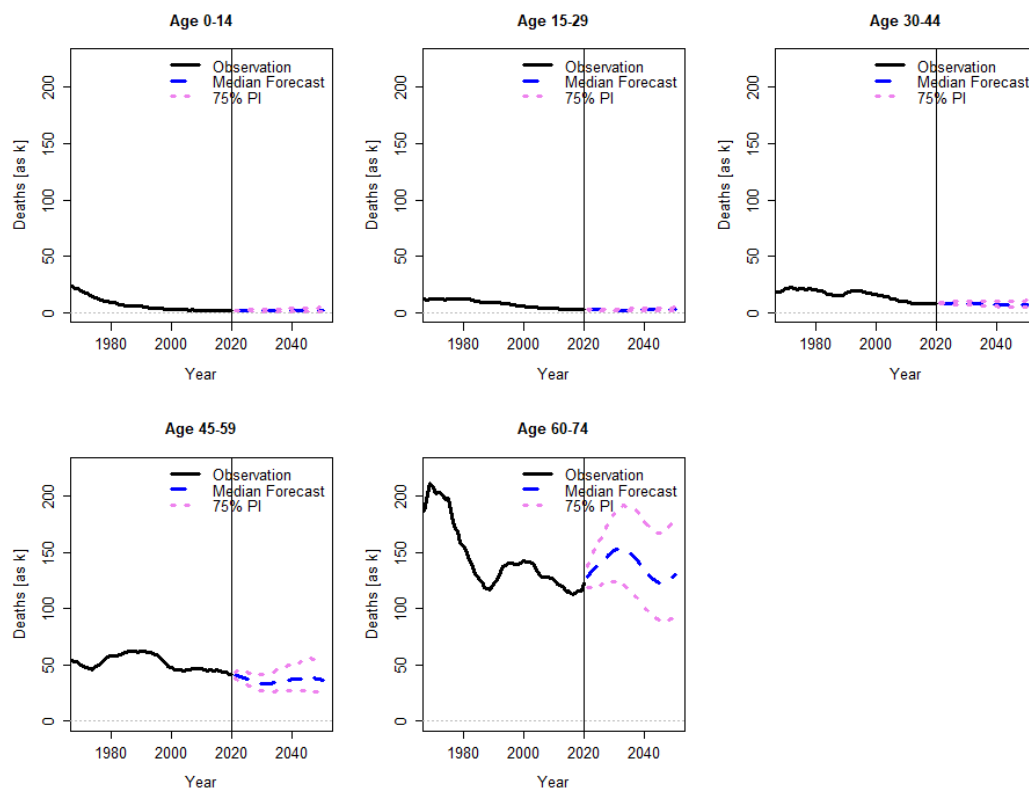
Figure A5. Forecasts of Deaths of Males by 15-Year Age Groups with 75% Prediction Intervals (Sources: [29,30]; authors' computation and illustration)
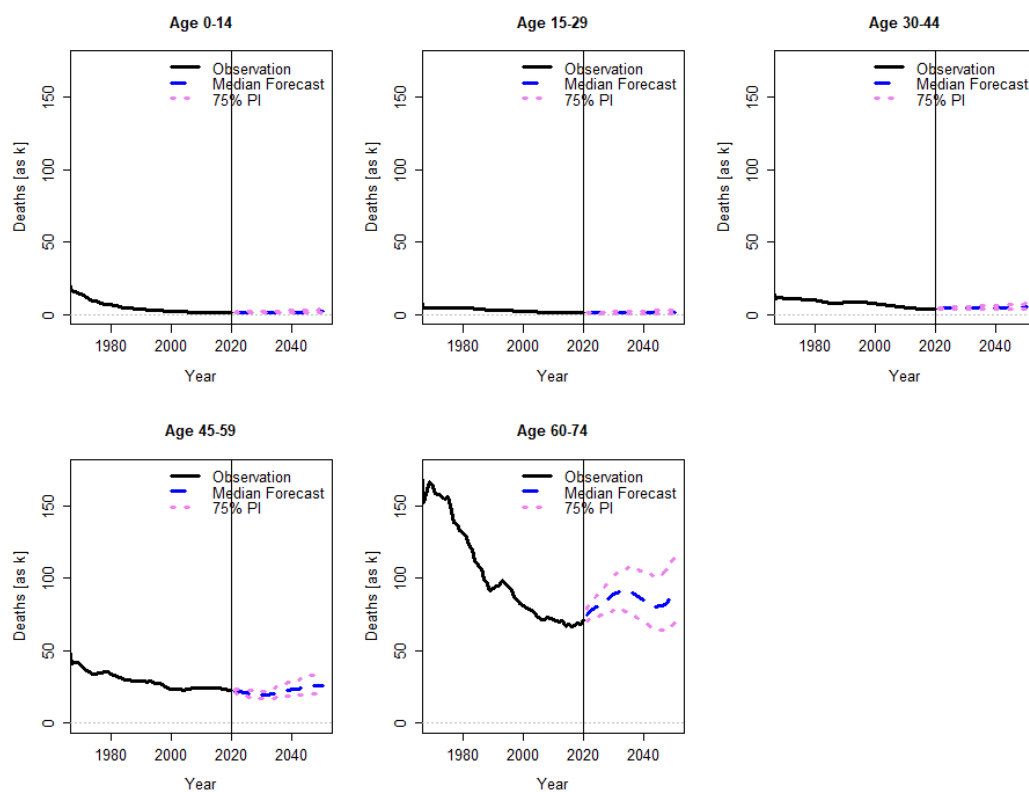

Figure A6. Forecasts of Deaths of Females by 15-Year Age Groups with 75% Prediction Intervals (Sources: [29,30]; authors' computation and illustration)

References

1.   Vanella, P.; Deschermeier, P. A Probabilistic Cohort-Component Model for Population Forecasting – The Case of Germany. *Journal of Population Ageing* **2020**, *13*, 513-545, doi:10.1007/s12062-019-09258-2.

2.   Wilke, C.B. Demografischer Wandel in Deutschland – Hintergründe, Zukunftsszenarien und Arbeitsmarktpotenziale. In *Megatrends aus Sicht der Volkswirtschaftslehre: Demografischer Wandel – Globalisierung & Umwelt – Digitalisierun*, Rebeggiani, L., Wilke, C.B., Wohlmann, M., Eds. Springer Gabler: Wiesbaden, 2020; 10.1007/978-3-658-30129-3_1 (pp. 3-24).

3.   Fuchs, J.; Sohnlein, D.; Weber, B.; Weber, E. Stochastic Forecasting of Labor Supply and Population: An Integrated Model. *Popul Res Policy Rev* **2018**, *37*, 33-58, doi:10.1007/s11113-017-9451-3.

4.   Vanella, P.; Rodriguez Gonzalez, M.A.; Wilke, C.B. The Impact of Population Aging on the German Statutory Pension Insurance – A Probabilistic Approach. In *Stochastic Forecasting in Demography and Social Insurance*, Vanella, P., Ed. Gottfried Wilhelm Leibniz Universität: Hannover, 2020; 10.15488/9408 (pp. 160-197).

5.   European Union. *The 2021 Ageing Report: Economic & Budgetary Projections for the EU Member States (2019-2070)*; Publications Office of the European Union: Luxembourg, 2021; 10.2765/84455.

6.   Vanella, P.; Heß, M.; Wilke, C.B. A probabilistic projection of beneficiaries of long-term care insurance in Germany by severity of disability. *Quality & Quantity* **2020**, *54*, 943-974, doi:10.1007/s11135-020-00968-w.

7.   BMAS; BAuA. *Sicherheit und Gesundheit bei der Arbeit - Berichtsjahr 2019. Unfallverhütungsbericht Arbeit*; BMAS: Paderborn, 2020; 10.21934/baua:bericht20201215.

8.   Vanella, P.; Deschermeier, P. A Principal Component Simulation of Age-Specific Fertility – Impacts of Family and Social Policy on Reproductive Behavior in Germany. *Population Review* **2019**, *58*, doi:10.1353/prv.2019.0002.

9.   Vanella, P. A principal component model for forecasting age- and sex-specific survival probabilities in Western Europe. *Zeitschrift für die gesamte Versicherungswissenschaft* **2017**, *106*, 539-554, doi:10.1007/s12297-017-0393-y.

10.   Wilke, C.B. Volkswirtschaftliche Kosten von Fehlzeiten in einer alternden Gesellschaft. In *Zur Relevanz von Bevölkerungsvorausberechnungen für Arbeitsmarkt-, Bildungs- und Regionalpolitik*, Deschermeier, P., Fuchs, J., Iwanow, I., Wilke, C.B., Eds. wbv: Bielefeld, 2020; 10.3278/301043w (pp. 118-139).

11.   Fuchs, J.; Söhnlein, D.; Vanella, P. Migration Forecasting—Significance and Approaches. *Encyclopedia* **2021**, *1*, 689-709, doi:10.3390/encyclopedia1030054.

12.   BAuA. *Arbeitswelt im Wandel: Zahlen - Daten - Fakten*; BAuA, Ed. Dortmund, 2019; 10.21934/baua:praxis20190131.

13.   Fuchs, J.; Söhnlein, D.; Weber, B. Projektion des Erwerbspersonenpotenzials bis 2060: Demografische Entwicklung lässt das Arbeitskräfteangebot stark schrumpfen. *IAB-Kurzbericht* **2021**, *25/2021*, URL:https://doku.iab.de/kurzber/2021/kb2021-25.pdf.

14.   Vanella, P.; Hellwagner, T.; Deschermeier, P. Parsimonious Stochastic Forecasting of International and Internal Migration on the NUTS-3 level – An Outlook of Regional Depopulation Trends in Germany. In Proceedings of Wittgenstein Centre Conference 2021, Vienna, Austria, 01 December 2021. URL:https://www.youtube.com/watch?v=QnLHKK4B0EQ.

15.   Destatis. Lebendgeborene nach dem Alter der Mütter - Insgesamt Alte Bundesländer 1960-1990. Provided on request on 18 June 2015.

16.   Destatis. Lebendgeborene nach dem Alter der Mütter - Insgesamt Alte Bundesländer 1961-1989. Provided on request on 9 February 2018.

17.   Destatis. Lebendgeborene nach dem Alter der Mütter - Neue Länder und Berlin-Ost. Berichtsjahr: 1990-2000. Provided on request on 13 February 2018.

18.   Destatis. Lebendgeborene nach dem Alter der Mütter - Insgesamt Deutschland 1991-2013. Provided on request on 18 June 2015.

19.   Destatis. Lebendgeborene: Deutschland, Jahre, Alter der Mutter, Geschlecht der Lebendgeborenen, Familienstand der Eltern. Available online: https://www-genesis.destatis.de/genesis/ (accessed on 13 December 2021).

20. Human Mortality Database. East Germany, Population size (abridged). Available online: www.mortality.org (accessed on 17 December 2021).

21. Human Mortality Database. Germany, Population size (abridged). Available online: www.mortality.org (accessed on 17 December 2021).

22. Human Mortality Database. West Germany, Population size (abridged). Available online: www.mortality.org (accessed on 17 December 2021).

23. Klüsener, S.; Grigoriev, P.; Scholz, R.D.; Jdanov, D.A. Adjusting Inter-censal Population Estimates for Germany 1987-2011: Approaches and Impact on Demographic Indicators. *Comparative Population Studies* **2018**, *43*, 31-64, doi:10.12765/CPoS-2018-06en.

24. Destatis. Bevölkerung am 31.12.2018-2020 nach Alters- und Geburtsjahren. Provided on request on 21 December 2021.

25. Destatis. Lebendgeborene: Deutschland, Jahre, Geschlecht. Available online: https://www-genesis.destatis.de/genesis/ (accessed on 13 December 2021).

26. Vanella, P.; Deschermeier, P.; Greil, A.L. Impacts of the COVID-19 Pandemic on International Fertility - A Stochastic Principal Component Approach. In Proceedings of Pandemic Babies? The Covid-19 Pandemic and Its Impact on Fertility and Family Dynamics, Rostock, 14 December 2021. Available at https://www.demogr.mpg.de/mediacms/16382_main_MPI2021_PV.pdf.

27. Destatis. Wanderungen zwischen Deutschland und dem Ausland 1991-1999 nach Einzelaltersjahren und Geschlecht. Provided on request on 02 September 2015.

28. Destatis. Wanderungen zwischen Deutschland und dem Ausland: Deutschland, Jahre, Nationalität, Geschlecht, Altersjahre. Available online: https://www-genesis.destatis.de/genesis/ (accessed on 13 December 2021).

29. Destatis. Bevölkerung und Erwerbstätigkeit. Gestorbene nach Alters- und Geburtsjahren sowie Familienstand 1948 bis 2003. Provided on request on 18 March 2016.

30. Destatis. Gestorbene 2000-2020 nach Alters- und Geburtsjahren. Provided on request on 15 August 2017 - 20 December 2021.

31. Fritsch, F.N.; Carlson, R.E. Monotone Piecewise Cubic Interpolation. *SIAM Journal of Numerical Analysis* **1980**, *17*, 238-246, doi:https://doi.org/10.1137/0717021.

32. Kirk, A.; Stamatiadis, N. Crash Rates and Traffic Maneuvers of Younger Drivers. *Transportation Research Record: Journal of the Transportation Research Board* **2001**, *1779*, 68-74, doi:https://doi.org/10.3141/1779-10.

33. Beznoska, M.; Pimpertz, J.; Stockhausen, M. Führt eine Bürgerversicherung zu mehr Solidarität? Eine Vermessung des Solidaritätsprinzips in der gesetzlichen Krankenversicherung. *IW-Analysen* **2021**, *143*, URL:https://www.iwkoeln.de/studien/martin-beznoska-jochen-pimpertz-maximilian-stockhausen-eine-vermessung-des-solidaritaetsprinzips-in-der-gesetzlichen-krankenversicherung.html.

34. Leinert, J. Morbidität als Selektionskriterium. In *Fairer Wettbewerb oder Risikoselektion? Analysen zur gesetzlichen und privaten Krankenversicherung*, Jacobs, K., Klauber, J., Leinert, J., Eds. Wissenschaftliches Institut der AOK: Bonn, 2006;   (pp. 67-76).

35. Schulenburg, J.M.; Lohse, U. *Versicherungsökonomik: Ein Leitfaden für Studium und Praxis*; Verlag Versicherungswirtschaft: Karlsruhe, 2014; Vol. 2.

36. Stauder, J.; Kossow, T. Selektion oder bessere Leistungen – Warum sind Privatversicherte gesünder als gesetzlich Versicherte? *Gesundheitswesen* **2017**, *79*, 181-187, doi:10.1055/s-0042-104583.

37. Singh-Manoux, A.; Gueguen, A.; Ferrie, J.; Shipley, M.; Martikainen, P.; Bonenfant, S.; Goldberg, M.; Marmot, M. Gender differences in the association between morbidity and mortality among middle-aged men and women. *Am J Public Health* **2008**, *98*, 2251-2257, doi:10.2105/AJPH.2006.107912.

38. Destatis. Studienanfängerquote (Hochschulzugangsberechtigung): Deutschland, Jahre, Geschlecht. Available online: https://www-genesis.destatis.de/genesis/ (accessed on 09 November 2021).

39. Vanella, P.; Deschermeier, P.; Wilke, C.B. An Overview of Population Projections—Methodological Concepts, International Data Availability, and Use Cases. *Forecasting* **2020**, *2*, 346-363, doi:10.3390/forecast2030019.

40.   Fries, J.F. Aging, Natural Death, and the Compression of Morbidity. *New England Journal of Medicine* **1980**, *303*, 130-135, doi:10.1056/NEJM198007173030304.

41.   Gruenberg, E.M. The Failures of Success. *Milbank Quarterly* **2005**, *83*, 779-800, doi:10.1056/NEJM198007173030304.

42.   Vanella, P.; Basellini, U.; Lange, B. Assessing excess mortality in times of pandemics based on principal component analysis of weekly mortality data-the case of COVID-19. *Genus* **2021**, *77*, 16, doi:10.1186/s41118-021-00123-9.

43.   Vanella, P.; Wiessner, C.; Holz, A.; Krause, G.; Möhl, A.; Wiegel, S.; Lange, B.; Becher, H. Pitfalls and solutions in case fatality risk estimation – A multi-country analysis on the effects of demographics, surveillance, time lags between case reports and deaths and healthcare system capacity on COVID-19 CFR estimates. *Vienna Yearbook of Population Research* **2022**, *20* [forthcoming], doi:https://doi.org/10.1553/populationyearbook2022.res1.4.

44.   Klinger, S.; Fuchs, J. Country report: GERMANY. In *Effects of population changes in the labour market: an analysis of six European countries*, Räisänen, H., Maunu, T., Eds. Ministry of Economic Affairs and Employment: Helsinki, 2019; Vol. 2019:59,   (pp. 54-66).

45.   Khailaie, S.; Mitra, T.; Bandyopadhyay, A.; Schips, M.; Mascheroni, P.; Vanella, P.; Lange, B.; Binder, S.C.; Meyer-Hermann, M. Development of the reproduction number from coronavirus SARS-CoV-2 case data in Germany and implications for political measures. *BMC Med* **2021**, *19*, 32, doi:10.1186/s12916-020-01884-4.

46.   Bauer, A.; Weber, E. COVID-19: how much unemployment was caused by the shutdown in Germany? *Applied Economics Letters* **2020**, *28*, 1053-1058, doi:10.1080/13504851.2020.1789544.