On the Impact of Separate Fertility and Mortality Assumptions for Native and Migrant Subpopulations in Population Projections

Christina Bohk

March 1, 2011

Abstract

Several demographic studies have shown that fertility and mortality often differ between natives and migrants. Hence, to accurately project a population on a macro level, a population should be divided into such subpopulations that do not differ substantially in their demographic behavior. However, common population projections often neglect this aspect due to the lack of data. Yet, if data are available, a single projection of natives, immigrants, emigrants, and their descendant generations with specific fertility and mortality assumptions reduces the effect of demographic heterogeneity and increases a population projection's accuracy. The impact of separate assumptions for native and migrant subpopulations is evaluated by comparing four projections, based on real-world data for Germany, using the probabilistic population projection model (PPPM). It turns out that assuming equivalent mortality and fertility for natives and migrants induces a considerable error in long-term projections, for both gross and net migration.

1 Introduction

Accurate population projections are indispensable for several societal and political decision making processes. For instance, the future size and structure of a population by age and sex are relevant issues to pre-plan the health care or pension system properly.

2 Background and Related Work

Several demographic studies have shown that fertility and mortality often differ between natives and migrants [1]. In Germany, the higher fertility and lower mortality of immigrants approaches the fertility and mortality level of the natives over time [2, 3, 4, 5, 6, 7]. This can be modeled by adjusting fertility and mortality levels for each subsequent descendant generation of immigrants. Hence, to accurately project a population on a macro level, separate fertility and mortality assumptions should be considered for native and migrant subpopulations.

Common population projection models, however, often assume equivalent fertility and mortality for immigrants and natives [8, 9, 10, 11]. Such models can neither project the actual size and structure of the immigrants, nor can they project an adaption between native and migrant fertility and mortality (over successive descendant generations). This simplification is often motivated by a lack of data — but for some countries (e.g., Germany) the fertility and mortality of the foreign population can be used to estimate the missing data. To warrant the efforts required for such estimations, it has to be shown that they improve the accuracy of a population projection.

The probabilistic population projection model (PPPM) avoids the aforementioned simplification. It allows for a single projection of natives, immigrants, emigrants, and their descendant generations with *separate* fertility and mortality assumptions.[1] Therefore, the PPPM lends itself particularly well to analyze the impact of separate fertility and mortality assumptions on several subpopulations.

The PPPM is a probabilistic macro model to project a population by age and sex. Besides supporting separate assumptions for different subpopulations, the PPPM has some additional advantages compared to common approaches. One advantage is the unrestricted generation of assumption paths[2] without a pre-determined method. Another advantage is the subjective assignment of an expected likelihood (i.e., occurrence probability) to each assumption path. Together, these two features enable a flexible construction of any assumption distribution a forecaster might expect. This flexibility pertains all model parameters. It allows to closely reflect various sources of information, e.g. past development, theories, or surveys.

Finally, the PPPM allows to account for correlations among assumption paths and offers different modes of simulative and (quasi-)exhaustive computation. Its current implementation is based on the modeling and simulation framework JAMES II [18, 19].

3 Data and Method

To evaluate the projection error of net migration with equivalent assumptions for native and migrant subpopulations — compared to gross migration and separate assumptions for native and migrant subpopulations — the PPPM is used to conduct four population projections, based on real-world data for Germany. All four projections project the total population (subdivided into natives, immigrants, emigrants, and three descendant generations each) by single age (up to $95+$) and sex over the projection horizon from 2007 to 2050.

The projections *Gross Migration I* and *Gross Migration II* consider both gross migration, but the former regards separate and the latter equivalent assumptions on fertility and mortality for all subpopulations. Analogously, the projections *Net Migration I* and *Net Migration II* consider both net migration, but the former regards separate and the latter equivalent assumptions on fertility and mortality for all subpopulations (see table 1).

[1] The PPPM is an extension of several basic approaches that divide a population into native and migrant subpopulations [12, 13, 14, 15, 16, 17].

[2] An assumption path contains the assumed future values of a model parameter over the whole projection horizon.

	Gross migration	**Net migration**
Separate assumptions	*Gross Migration I*	*Net Migration I*
Equivalent assumptions	*Gross Migration II*	*Net Migration II*

Table 1: Characteristics of the four population projections regarding gross/net migration and separate/equivalent assumptions for all subpopulations

To evaluate the impact due to separate assumptions, all gross migration assumption paths are transformed into corresponding net migration assumption paths. The assumptions on fertility and mortality are the same for all subpopulations, except for the immigrants and their descendant generations in the projections *Gross Migration I* and *Net Migration I*. For the latter two projections, the higher fertility and lower mortality of immigrants are assumed to adjust successively, over the descendant generations, to the level of the natives (see table 3).

Each projection considers several assumption paths for a total of 87 model parameters (with three descendant generations for immigrants and emigrants); a list of these 87 model parameters is given in table 2.[3]

			$ASFR^{*}$	l_x^{**}	$\pi(P_0)^{m***}$	$\pi_{(0,5}D_0)^{****}$	p_{95+}^{*****}	Total numbers
Natives		Female	•	•	•	•	•	$P_{x,31.12.Jump-off}^{******}$
Natives		Male		•	•	•	•	$P_{x,31.12.Jump-off}$
Immigrants	Direct Immigrants	Female	•	•	•	•	•	$Z_x^{*******}$
Immigrants	Direct Immigrants	Male		•	•	•	•	Z_x
Immigrants	Desc. Gen. 1	Female	•	•	•	•	•	
Immigrants	Desc. Gen. 1	Male		•	•	•	•	
Immigrants	Desc. Gen. 2	Female	•	•	•	•	•	
Immigrants	Desc. Gen. 2	Male		•	•	•	•	
Immigrants	Desc. Gen. 3	Female	•	•	•	•	•	
Immigrants	Desc. Gen. 3	Male		•	•	•	•	
Emigrants	Direct Emigrants	Female	•	•	•	•	•	$Fz_x^{********}$
Emigrants	Direct Emigrants	Male		•	•	•	•	Fz_x
Emigrants	Desc. Gen. 1	Female	•	•	•	•	•	
Emigrants	Desc. Gen. 1	Male		•	•	•	•	
Emigrants	Desc. Gen. 2	Female	•	•	•	•	•	
Emigrants	Desc. Gen. 2	Male		•	•	•	•	
Emigrants	Desc. Gen. 3	Female	•	•	•	•	•	
Emigrants	Desc. Gen. 3	Male		•	•	•	•	

* ASFR — age-specific fertility rates
** l_x — survivors at age x
*** $\pi(P_0)^m$ — proportion of male births
**** $\pi_{(0,5}D_0)$ — share of infant deaths that die in first half year
***** p_{95+} — survival probability of persons aged 95 and above
****** $P_{x,31.12.Jump-off}$ — Year-end jump-off population by age
******* Z_x — Total numbers of direct immigrants by age
******** Fz_x — Total numbers of direct emigrants by age

Table 2: List of 87 model parameters in each population projection

For each subpopulation, there are four assumption paths regarding fertility[4] and five to six assumption paths regarding mortality.[5] Each assumption path is generated by a combination of methods, e.g., time series analysis and corrective expert judgment (based on theory). To capture the uncertainty with respect to a model parameter, its assumption paths vary in level, pattern, and expected likelihood. Table 3 shows the

[3] In Bohk et al. [18] is also a figure that gives an overwiew of the model parameters and subpopulations in the PPPM.

[4] More precisely, age-specific fertility rates over the projection horizon.

[5] More precisely, sex-specific survivors at age x, over the projection horizon.

mortality and fertility assumption paths with the highest expected likelihood for all subpopulations represented by Total Fertility Rate (TFR) and life expectancy at birth (e_0) in 2050 for the two projections *Gross Migration I* and *Net Migration I*.[6]

			TFR	pr_{TFR}	e_0	pr_{e_0}
Natives		Female	1.42	0.6	88.41	0.3
Natives		Male			85.32	0.3
Immigrants	Direct Immigrants	Female	1.84	0.5	92.63	0.4
Immigrants	Direct Immigrants	Male			91.68	0.4
Immigrants	Desc. Gen. 1	Female	1.69	0.5	89.59	0.45
Immigrants	Desc. Gen. 1	Male			87.22	0.45
Immigrants	Desc. Gen. 2	Female	1.46	0.75	88.41	0.4
Immigrants	Desc. Gen. 2	Male			85.32	0.4
Immigrants	Desc. Gen. 3	Female	1.42	0.6	88.41	0.3
Immigrants	Desc. Gen. 3	Male			85.32	0.3
Emigrants	Direct Emigrants	Female	1.42	0.6	88.41	0.3
Emigrants	Direct Emigrants	Male			85.32	0.3
Emigrants	Desc. Gen. 1	Female	1.42	0.6	88.41	0.3
Emigrants	Desc. Gen. 1	Male			85.32	0.3
Emigrants	Desc. Gen. 2	Female	1.42	0.6	88.41	0.3
Emigrants	Desc. Gen. 2	Male			85.32	0.3
Emigrants	Desc. Gen. 3	Female	1.42	0.6	88.41	0.3
Emigrants	Desc. Gen. 3	Male			85.32	0.3

Table 3: Mortality and fertility assumption paths with highest expected likelihood (pr) for all subpopulations in the two projections *Gross Migration I* and *Net Migration I*; Total Fertility Rate (TFR) and life expectency at birth (e_0) in 2050 represent the corresponding fertility and mortality assumption paths

Despite a realistic setting, the purpose of the four projections is to evaluate the impact of separate assumptions for native and migrant subpopulations. Consequently, they should not be taken for real forecasts.

4 Results

The results of each projection are computed by 6.000 trials of a Monte-Carlo simulation. As the impact of separate assumptions and the consideration of migration (net vs. gross) shall be observed, only result *differences* are reported in the following. Differences are calculated by subtracting the projected median results: the median results in case of separate assumptions (*Net/Gross Migration I*) are subtracted by the corresponding median results in case of equivalent assumptions (*Net/Gross Migration II*). The differences are plotted in figure 1.

The differences between *Gross Migration I* and *Gross Migration II* (top row, fig. 1)

[6] Additional information are also available upon request.

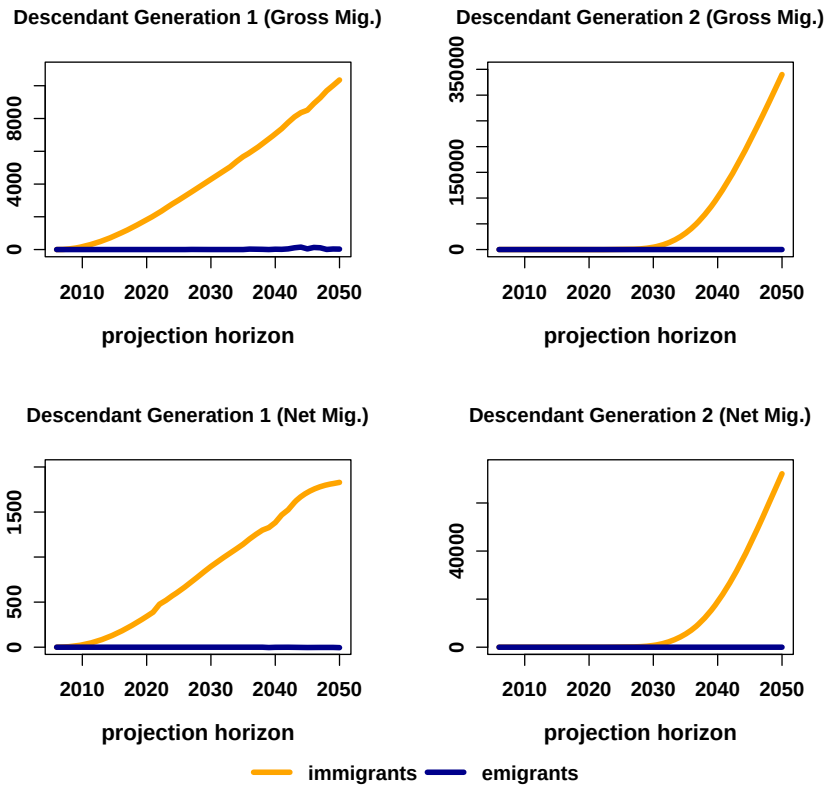

Figure 1: Population difference between *Gross Migration I* and *II* (top) as well as *Net Migration I* and *II* (bottom). Differences are given for the first (left) and the second (right) descendant generation of immigrants (orange) and emigrants (blue).

are substantial for immigrants,[7] whereas they are negligible for emigrants. Furthermore, the difference between immigrant descendants increases with time and generation. Similar observations can be made for net migration (bottom, fig. 1): the overall pattern is the same as for gross migration, but on a much smaller scale.

The large differences for immigrants clearly stem from having separate assumptions on fertility and mortality. In contrast, results are almost identical for the emigrants: no separate assumptions are considered, so differences are due to stochastic noise. Since the separate assumptions for immigrants account for their lower mortality and higher fertility, they induce considerably more offspring in the projections *Gross/Net Migration I* than in their counterparts with equivalent assumptions. The underestimated offspring of the immigrants in the projections *Gross/Net Migration II* can be regarded as projection error. As each descendant generation emanates from its former generation, this projection error — once occurred — accumulates in each

[7] In 2050, the second descendant generation of *Gross Migration I* is about 30% larger than in *Gross Migration II*.

succeeding generation. Hence, this projection error increases as (1) the number of immigrants increases and (2) the difference between immigrants' and all other (native and emigrant) subpopulations' fertility and mortality assumptions increases.[8]

5 Conclusions

This work investigates the impact of separate assumptions for native and migrant subpopulations in a population projection. Four probabilistic population projections, based on real-world data from Germany, have been conducted with the PPPM. A comparison of the results shows a loss of accuracy when assuming equivalent (instead of separate) fertility and mortality for migrants and natives. Unsurprisingly, the errors due to this simplification may cumulate (see fig. 1) from generation to generation — which could be particularly harmful for the validity of long-term projections. The results also show that the effect is unlikely to invalidate short-term projections.

[8] Given a higher fertility and lower mortality for the immigrants than for the other (native and emigrant) subpopulations.

References

[1] Tomas Sobotka. The rising importance of migrants for childbearing in Europe. *Demographic Research*, 19:225–248, 2008.

[2] Nadja Milewski. First child of immigrant workers and their descendants in West Germany: Interrelation of events, disruption, or adaption? *Demographic Research*, 17(29):859–896, 2007.

[3] Nadja Milewski. *Fertility of immigrants and their descendants in West Germany: an event-history approach*. PhD thesis, University of Rostock, 2008.

[4] Nadja Milewski. Immigrant fertility in West Germany: Is there a socialisation effect in transitions to second and third births? *European Journal of Population*, 26:297–323, 2010.

[5] Martin Kohls. Healthy-Migrant-Effect, Erfassungsfehler und andere Schwierigkeiten bei der Analyse der Mortalität von Migranten. *Bundesamt für Migration und Flüchtlinge, Working Paper 15*, 2008.

[6] Martin Kohls. Leben Migranten wirklich länger? Eine empirische Analyse der Mortalität von Migranten in Deutschland. *Bundesamt für Migration und Flüchtlinge, Working Paper 16*, 2008.

[7] Martin Kohls. Mortality of Immigrants in Germany. In *Proceedings of the European Population Conference 2008*, 2008.

[8] Wolfgang Lutz, Warren C. Sanderson, and Sergei Scherbov. Expert- based probabilistic population projections. *Population and Development Review, Supplement: Frontiers of Population Forecasting*, 24:139–155, 1998.

[9] Wolfgang Lutz, Warren C. Sanderson, and Sergei Scherbov. The end of world population growth. *Nature*, 412:543–545, August 2001.

[10] Juha M. Alho and J. Nyblom. Mixed estimation of old-age mortality. *Mathematical Population Studies*, 6(4):319–330, 1997.

[11] Juha M. Alho and Bruce D. Spencer. *Statistical Demography and Forecasting*. Springer Science+Business Media, Inc., 2005.

[12] Thomas J. Espenshade, Leon F. Bouvier, and W. Brian Arthur. Immigration and the stable population model. *Demography*, 19:125–133, 1982.

[13] W. Brian Arthur and Thomas J. Espenshade. Immigration Policy and Immigrants' Ages. *Population and Development Review*, 14(2):315–326, 1988.

[14] Barry Edmonston and Jeffrey S. Passel. Immigration and immigrant generations in population projections. *International Journal of Forecasting*, 8:459–476, 1992.

[15] Samarendranath Mitra. Generalization of immigration and the stable population model. *Demography*, 20:111–115, 1983.

[16] Pietro Cerone. On stable population theory with immigration. *Demography*, 24:431–438, 1987.

[17] Carl P. Schmertmann. Immigrant's ages and the structure of stationary populations with below-replacement fertility. *Demography*, 29:595–612, 1992.

[18] Christina Bohk, Roland Ewald, and Adelinde M. Uhrmacher. Probabilistic Population Projection with JAMES II. In M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, editors, *Proceedings of the Winter Simulation Conference*, 2009.

[19] Thomas Salzmann and Christina Bohk. The unexpected results produced by a specific design of a probabilistic population projection model. In *Proceedings of the Population Association of America 2006 Annual Meeting*, Los Angeles, March 2006.