Correlatedness of Mortality Assumptions in a Probabilistic Population Projection Model

Christina Bohk* and Roland Rau*

*University of Rostock

1 Introduction

In the public, demography gains special attention with population forecasts. This is especially due to their great societal importance regarding, e. g., political planning of the social security system.

However, according to Keilman [1], European population projections' accuracy did not improve substantially so far. Therefore, the technique of population forecasting should be further developed in order to reduce projection error. In this presentation, we address an important methodological issue in probabilistic population forecasting: the correlatedness of demographic events, and its impact on the accuracy of projection outcome.

2 Correlatedness of demographic events

It is often discussed in demography, if mortality, fertility, and migration are *inter*-dependent, or if their development is *self*-dependent, and just an autonomous reaction to the same environment [2, 3]. An example for interdependence is, for instance, the first demographic transition. According to this theory, the decline in mortality triggers to a large extent the decline in fertility, although contrary observations have been made in many countries [4, 5, 6]. The second demographic transition [7] could provide a counterexample of *interdependence*: an increase in fertility did go along with a decline in mortality due to different reasons in many industrialized countries.

Next to the *inter*- or *in*-dependence of different vital events, it is also an interesting research question if there is any kind of *intra*-dependence of a vital event like mortality. Many studies in demography show that mortality differs among subpopulations [8]. For Germany, they indicate that immigrants' mortality is lower than mortality of the autochthonous population (due to a healthy-migrant effect), and that immigrants' mortality approaches the mortality level of the autochthonous population over successive descendant generations [9, 10]. Similar results have also been found in Germany for fertility, i. e., that immigrants higher fertility approaches lower fertility of the autochthonous population over successive descendant generations [11, 12]. Disregarding such demographic heterogeneity in mortality (and/or fertility) can induce considerable projection error [13].

3 Relevance for population projections

The question of correlatedness within and among demographic events is highly relevant in probabilistic projections. A typical simulative procedure combines randomly chosen assumptions for mortality, fertility, and migration in each trial. But which of these multiple assumptions can be combined plausibly, and even more important, which of these multiple assumptions cannot be combined plausibly at all? And if the combination of some assumptions is actually implausible, to what extent can they deteriorate projection outcome?

4 Application in a population projection for Germany

To analyze the impact of correlatedness within and among assumptions for vital rates on projection outcome, we conduct two probabilistic projections for Germany up to 2050 with a novel framework for probabilistic population forecasting, entitled PPPM.[1] The PPPM is a probabilistic model that projects the autochthonous population, immigrants and their descendant generations with separate assumptions for vital events. Therefore, it allows a forecaster to better incorporate demographic heterogeneity in mortality and fertility.

In this paper, we focus on mortality to explain the concept of how to set correlations among assumptions in the PPPM. In both projections—the one with and the other without set correlations—, we consider six assumptions for each subpopulation that vary in their mortality level, trajectory over age and time and expected likelihood of occurrence.

In the *first projection*, we do not set correlations among mortality assumptions for the different subpopulations, i. e. for the autochthonous population, the immigrants and their descendant generations. As a consequence, all mortality assumptions can be combined with each other in a projection trial.

In contrast, we do set correlations among mortality assumptions for the different subpopulations in the *second projection*. The purpose is to assure that our expectation, derived from the literature, that immigrants' lower mortality approaches higher mortality of the autochthonous population over descendant generations cannot be violated in any projection trial. As a consequence, we set correlations to (1) exclude implausible combinations of mortality assumptions, and to (2) increase projection outcomes' accuracy. We set these correlations in the PPPM with a new general framework, i. e. with *Settypes* and *Sets*. In a Settype, we comprise mortality for all subpopulations as model parameters. Afterwards, we create different Sets with specific mortality assumptions for each subpopulation for this given Settype. One Set contains only assumptions with relatively low mortality for all subpopulations, whereas all included assumptions for the direct immigrants have even lower mortality than those for the natives. Next to the Set with relatively low mortality, there are also Sets with medium and relatively high mortality. To conduct a projection trial, the PPPM randomly chooses a Set first, and a mortality assumption for each subpopulation from this chosen Set thereafter.

Figure 1 depicts for both projections the cumulative distribution function for the total population in 2050. The total population will lie with a probability of 90 percent between 62.3 and 77.4

[1] An implementation of the PPPM is open source and freely available at: `https://bitbucket.org/Christina_Bohk/p3j/wiki/Home`.

million according to the first projection, and between 63.7 and 73.5 million according to the second projection. Obviously, the variance of the projection outcome is larger in the first than in the second projection. This is mainly due to the reduction of possible assumption combinations via Settypes and Sets in the second projection. Hence, in this example, implausible assumption combinations induce extreme low as well as extreme high total population counts.

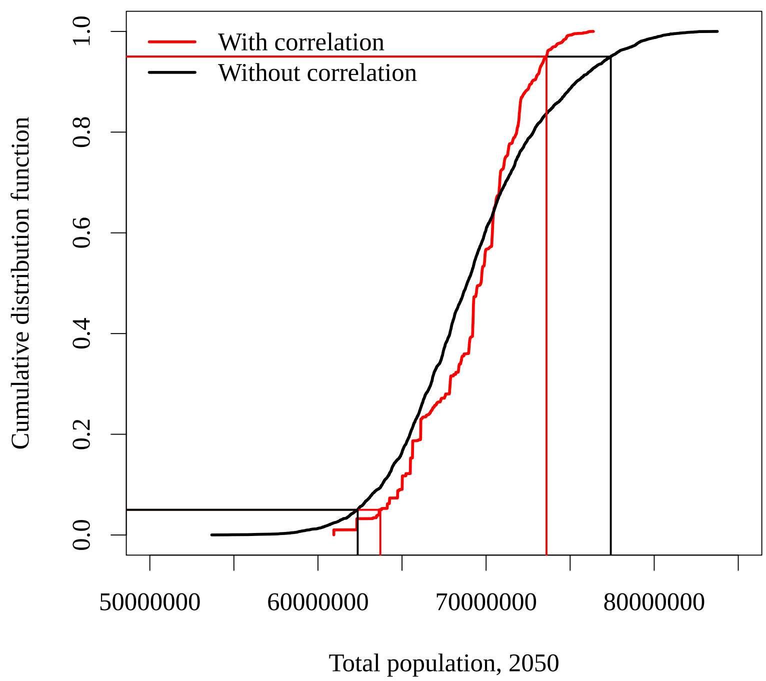

Figure 1: Cumulative distribution function for total population of Germany in 2050.

	With correlation			**Without correlation**		
	0.05	**0.5**	**0.95**	**0.05**	**0.5**	**0.95**
2020	79,273,089	80,249,150	80,697,371	79,233,096	80,036,488	80,897,200
2030	75,081,777	77,386,823	78,756,475	74,461,769	77,050,329	79,504,621
2040	69,775,827	73,796,677	76,280,118	69,135,958	73,231,516	78,197,366
2050	63,711,869	69,688,687	73,594,360	62,362,858	68,872,748	77,418,012

Table 1: Total population counts for Germany for the quantiles 0.05, 0.5, and 0.95 in the years 2020, 2030, 2040, and 2050 from the projection *with* and *without* correlation among mortality assumptions of different subpopulations.

5 Concluding remarks

The PPPM provides an innovative and simple way to set correlations within and among assumptions of vital rates with Settypes and Sets. This general framework allows us to intuitively exclude implausible combinations of assumptions for mortality, fertility, and migration. As a result, we can reduce projection error as well as increase projection outcomes' accuracy substantially. In this paper, we will provide an in-depth analysis of the impact of correlations among assumptions for vital events on projection outcome, with a focus on mortality.

References

[1] Nico Keilman. European Demographic Forecasts Have Not Become More Accurate Over the Past 25 Years. *Population and Development Review*, 34(1):137–153, 2008.

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

[3] Nico Keilman. Ex-post errors in official population forecasts in industrialized countries. *Journal of Official Statistics*, 13:245–277, 1997.

[4] F. W. Notestein. Population: the long view. In T. W. Schultz, editor, *Food for the world*, pages 36–57. University of Chicago Press, 1945.

[5] K. Davis. The world demographic transition. *The Annals of the American Academy of Political and Social Science*, 237:1–11, 1945.

[6] Henk A. de Gans. *Population Forecasting 1895-1945. The Transition to Modernity*. Kluwer Academic Publishers, 1999.

[7] Dirk J. van de Kaa. Europe's second demographic transition. *Population Bulletin*, 42(1):1–59, 1987.

[8] Laura Blue and Andrew Fenelon. Explaining low mortality among US immigrants relative to native-born Americans: the role of smoking. *International Journal of Epidemiology*, 40:786–793, 2011.

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

[10] Martin Kohls. *Demographie von Migranten in Deutschland*. Challenges in Public Health. Peter Lang, 2012.

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

[12] 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.

[13] Christina Bohk. On the Impact of Separate Fertility and Mortality Assumptions for Native and Migrant Subpopulations in Population Projections. In *PAA Annual Meeting 2011*, 2011.