Slovak University of Technology in Bratislava
Faculty of Mechanical Engineering
16th Conference on Applied Mathematics **APLIMAT 2017**
Proceedings

AN APPLICATION OF POLYNOMIAL FUNCTIONS OF DIFFERENT ORDER IN SMOOTHING OF MORTALITY CURVES: THE CASE STUDY OF BULGARIA AND CZECHIA FROM 1960 TO 2010

DOTLAČILOVÁ Petra (CZ), ŠIMPACH Ondřej (CZ)

Abstract. In this paper we use classical polynomial functions of 2nd, 3rd and 4th order for smoothing the mortality curves at advanced ages (60–110 years of life) in the case of Bulgaria and Czechia. Classical polynomial functions are good approach as a supplement to weighted moving averages and commonly used models for smoothing variance in the age-and-sex-specific death rates. Our results are compared with each other using adjusted index of determination and it will be shown that the quality of smoothing is at very good level in the case of Bulgaria and Czechia.

Key words: polynomial function, mortality curve, age-specific death rate, Bulgaria, Czechia

Mathematics Subject Classification: Primary 90C30; Secondary 62H12

1 Introduction

Mortality at the advanced ages (60 years and above) becomes important topic for demographers and for analysts in the field of health and pension insurance and also for public policy planning [15]. Extensions of length of human life is influenced mainly by improvement in medicine and in healthcare. For our future it means, that the proportion of persons at higher ages will continue to increase (see e.g. papers by Fiala [7], [8] or Gavrilov, Gavrilova [9]). One cause of population ageing is decrease in mortality. But the empirical values of age-specific death rates show fluctuations, which are noticeable mainly at higher ages. That is the reason why is useful to smooth these values. At higher ages (60 years and above), where mortality has more natural character and variability in observations is higher, it is possible to use some of available functions (e.g. Gompertz-Makeham, Kanistö model, Coale-Kisker model, Thatcher model and others). On the other hand classical mathematical approaches as e.g. weighted moving averages or polynomial functions of different order could be used for smoothing of death rates as well.

The countries of Central and Eastern Europe is aging as it is in the case of developed Western European countries [13], [14]. At first, there was a significant improvement in a care of live born persons and infants, which caused the decrease of infant mortality and mortality rates during childhood. Later, mortality began to improve even at the advanced ages. Among reasons for this evolution could be included higher level of health care, standard of living, better living conditions or functional health and pension and social system. Another equally important reason could be more interest in a healthy lifestyle and also better environment in most countries (especially in industrial cities). Given to this evolution, it is important to have the best imagination about how long in average will live not only the youngest persons, but also the oldest ones.

In this paper we review the existing forms of levelling functions, which are used in mortality curves smoothing. After we describe the methodology of polynomial functions of 2nd, 3rd and 4th order and apply this approach on mortality curves of Bulgaria and Czechia for males and females separately, for the period from 1960 to 2010. Our results are compared with each other using adjusted index of determination and it is shown that the quality of smoothing is at very good level. Last sections of this paper discuss the results and concludes remarks with future challenges.

2 Models review

From the past analysis, performed e.g. by Boleslawski, Tabeau [1] or Burcin, Tesárková, Šídlo [3] it is evident that the level of mortality of younger persons is different in comparison with the oldest ones. Therefore, it is necessary to correct estimates of mortality at the highest ages. For this correction are used various types of models. Nowadays we can use several existing models which are used for smoothing and for estimating of unknown parameters. It is possible to obtain them by using the professional demographical software (e.g. DeRaS, see Burcin, Hulíková Tesárková, Kománek [2], Statistica, or IBM SPSS Statistics). Among the most famous are included Coale-Kisker model (see e.g. Coale, Kisker [4])

$$m_{x,t} = e^{ax^2+bx+c}, \tag{1}$$

where $m_{x,t}$ are age-and-sex specific death rates and x is age. Obtained model corresponds with an exponential quadratic function, where a, b and c are parameters of model. Next one is Thatcher model (see e.g. Thatcher, Kanistö, Vaupel [16])

$$\mu_{x,t} = \frac{z}{1+z} + \gamma, \tag{2}$$

where $z = \alpha e^{\beta x}$, α, β, γ are parameters of model, x is age (and μ_x is the intensity of mortality at exact age x)[1]. The other one is Kannistö model (Thatcher, Kanistö, Vaupel [16])

$$\mu_{x,t} = \frac{e^{[\theta_0+\theta_1(x-80)]}}{1+e^{[\theta_0+\theta_1(x-80)]}}, \tag{3}$$

where θ_0 and θ_1 are unknown parameters. Kanistö model is a special case of the logistic function, where logit transformation of death rates is expressed by linear function of age. The oldest model (but still very frequently used) is the Gompertz-Makeham function (Gompertz [10], Makeham [12])

[1] For the intensity of mortality is valid a simplistic equation $\mu(x+1/2) \doteq m_x$ for $x = 1,2,3,\ldots,59$ (see e.g. [7]).

$$\mu_{x,t} = a + bc^x, \tag{4}$$

where a, b and c are parameters. All these models are suitable for the elimination of fluctuations in age-specific death rates and their subsequent extrapolation (more e.g. in Ekonomov, Yarigin [6] or Dotlačilová, Šimpach, Langhamrová [5]). Another option is application of polynomial functions.

In this paper we use the polynomial functions of 2nd, 3rd and 4th order and estimation of parameters of regression functions will be based on the age range x = 60–85 years. We evaluate the significance of models by adjusted indices of determination $adj.R^2$ and this information will be one of the most important about the suitability of models. The other information about suitability can be comparison of smoothed and extrapolated values with theoretical expectations (according to Gompertz [10] theory about the law of mortality in a graphical form prepared by 3D Bicubic splines in *Statgraphics Centurion XVI* software).

3 Materials and Methods

For the purposes of mortality analysis we use demographic data from the "Human Mortality Database" (HMD [11]) for Bulgaria and Czechia, with annual frequency of observations. Let us denote the polynomial function of 2nd, 3rd and 4th order as

$$m_{x,t} = \begin{cases} \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \varepsilon_t \\ \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \varepsilon_t \\ \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \beta_4 x_t^4 + \varepsilon_t \end{cases}, \tag{5}$$

where $m_{x,t}$ are the age-specific mortality rates calculated by ratio

$$m_{x,t} = \frac{D_{x,t}}{E_{x,t}}, \tag{6}$$

where $D_{x,t}$ is the number of died x-years old persons in calendar year t, $E_{x,t}$ is the exposure to risk of x-years old persons in calendar year t, which is (commonly) estimated as a number of mid-year population x-years old in year t, β_0, β_1, β_2, β_3 and β_4 are parameters of polynomial functions and ε_t should be a residual element with characteristics of white noise process, where $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2$, $cov(\varepsilon_t\,;\,\varepsilon_t') = 0$ and $\varepsilon_t \approx$ N distribution, x is age and used range is $x \in < 0\,;\,110 >$. (The estimation of parameters of polynomial functions will be based on the age range x = 60–85 years only), t is time where $t \in < 1960; 2010 >$ for the case of Bulgaria and Czechia. These polynomial functions we use for smoothing of age-and-sex specific death rates from 60 to 85 years. This interval is not obligatory, but there are some authors (e.g. Burcin, Tesárková, Šídlo [3]), who recommend it. For the advanced ages we perform the extrapolation and our recommendation for the end is x = 110.

4 Results

There are modelled the intensities of mortality using polynomial functions of the 2nd, 3rd and 4th order in the following part. We chose Bulgaria (as the population of Eastern Europe with higher levels of mortality) and Czechia (as the population of Central Europe with a lower mortality rates). The resulting outputs are fed sequentially presented using 3D Bicubic splines. In Appendix of this paper is presented the evaluation of obtained outputs using adjusted coefficients of determination.

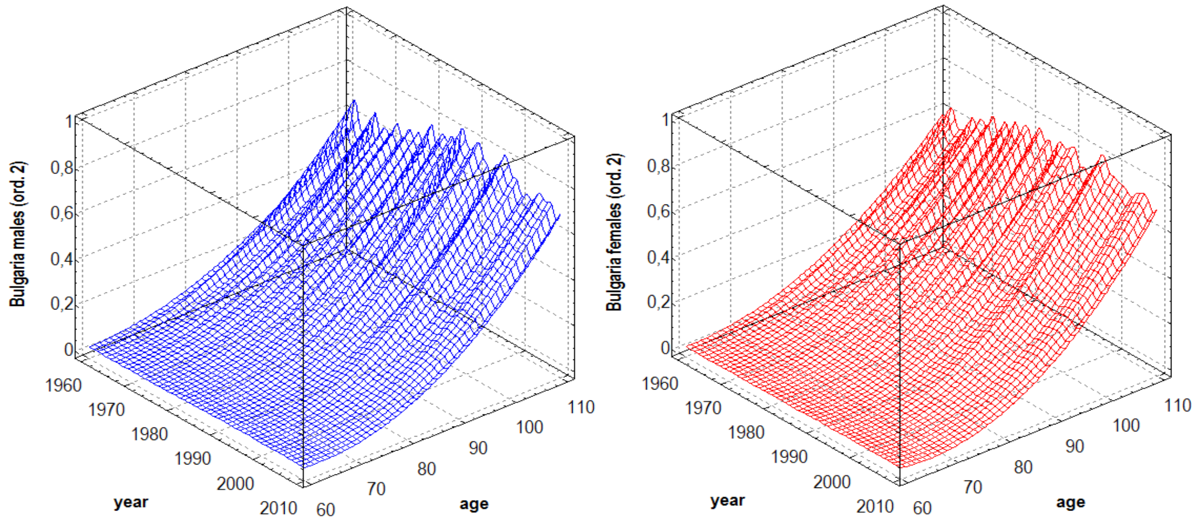

Fig. 1. Smoothing by polynomial function of the 2nd order for Bulgarian males (left) and females (right).
Source: authors' calculations and construction

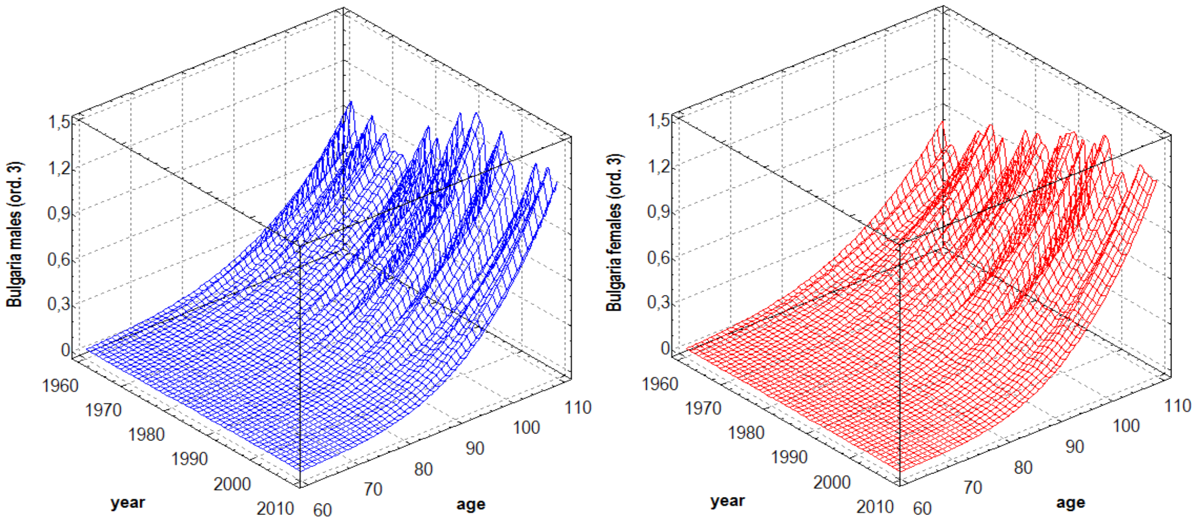

Fig. 2. Smoothing by polynomial function of the 3rd order for Bulgarian males (left) and females (right).
Source: authors' calculations and construction

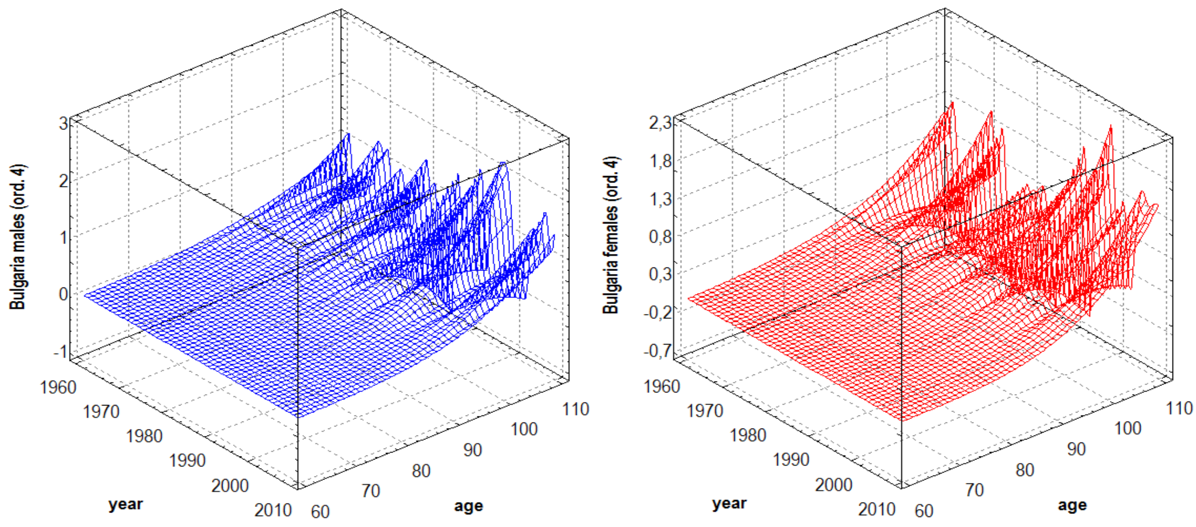

Fig. 3. Smoothing by polynomial function of the 4th order for Bulgarian males (left) and females (right).
Source: authors' calculations and construction

4.1 Bulgarian population

Fig. 1 shows the modelled intensity of mortality for the Bulgarian population using a polynomial function of the 2nd order. From the results it is evident that at high ages (90+) we receive relatively smooth fitting. This leads to significant removal of random and systematic variations, which are present in empirical death rates, especially in ages 85+. Fig. 2 shows modelled death rates for the same population using a polynomial function of the 3rd order. From the obtained values we can conclude that the selected function provides a higher level of mortality (in comparison with 2nd order of polynomial function). Then, it is also obvious that the function of 3rd order does not provide fluent fitting (especially in the highest ages). Fig. 3 shows modelled intensity of mortality using a polynomial of 4th order. When applying this polynomial for Bulgarian population we receive the least fluent fitting. (This is especially true in the highest ages – 100+). Moreover, it also shows that the model is not applicable for ages 100+, because there is a decrease in the mortality law (Gompertz [10]). Comparison of models using the adjusted indices of determination is presented in Fig. 7, 8 and 9 in Appendix of this paper. The best results are generally in the case of function of the 4th order, but that cannot be used universally. The significance of the polynomial of the 3rd order is sufficient in most cases and we can claim that can be used universally.

4.2 Czech population

Following part is devoted to modelling of mortality of the Czech population. Fig. 4 shows modelled intensity of mortality obtained when applying a polynomial function of the 2nd order. Also here is evident that we receive relatively smooth fitting in the highest ages. Moreover, modelled intensity of mortality corresponds to the assumption that the mortality decreases from past to present. In Fig. 5 is graphically shown the intensity of mortality modelled by polynomial function of 3rd order. If we compare the results obtained by function of the 2nd and 3rd order, we find that the second one provides higher values of the intensity of mortality. Then, it is also possible to say that we receive for Czech males less continuous fitting in the highest ages (compared to the previous function). Last

Fig. 6 shows the modelling of intensity of mortality using a polynomial function of 4th order. It is obvious even in the case of the Czech population that this function fails in some years. There is a reverse decline of intensity of mortality (breaking down the extrapolated development to the opposite direction). Compared to the previous functions is thus applicable in selected years only. Comparison of models for the Czech population using the adjusted indices of determination is presented in Fig. 10, 11 and 12 in Appendix of this paper. Again in this case are the best results generally provided by function of the 4th order. Using this function is not possible in general. The significance of the polynomial of the 3rd order was sufficient in most cases during development of 1960–2010 and then we can claim that it can be universally used.

5 Discussion and Conclusion

The aim of this paper was to show different ways of smoothing mortality curve at ages 60+. As an analytical functions were chosen polynomial functions of the 2nd, 3rd and 4th order. By these functions there was modelled the intensity of mortality in Bulgarian and Czech population. Another goal was to assess the obtained fitting using by *adj.*R^2. Based on the results we can say that using a polynomial function of the 2nd and 3rd order brings relatively smooth fitting. A different situation arises when applying a function of the 4th order. This function is failing mainly in the case of Bulgarian population, and this misalignment is evident mainly in the highest ages (about 100 years).

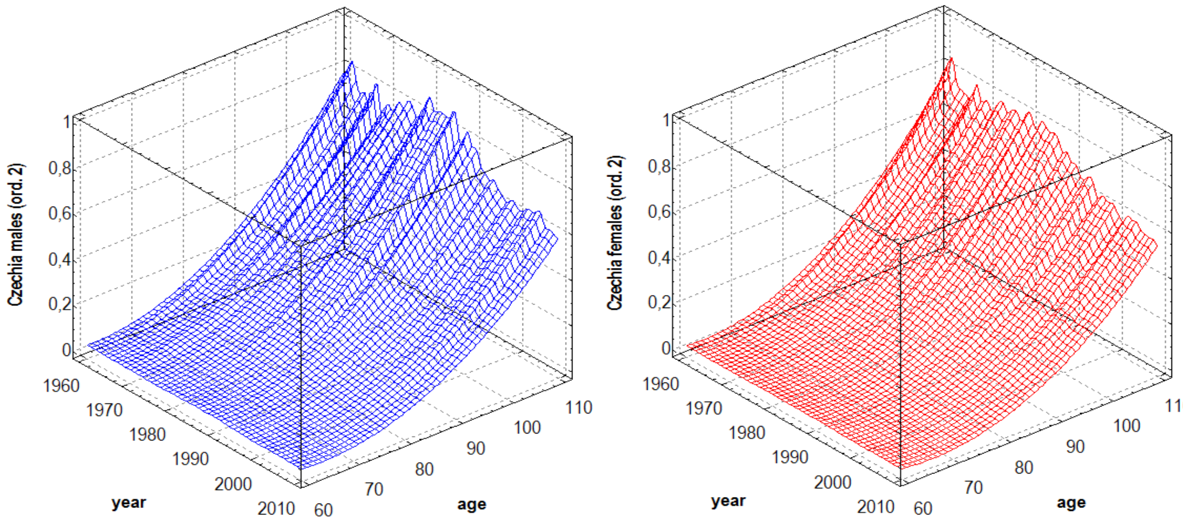

Fig. 4. Smoothing by polynomial function of the 2nd order for Czech males (left) and females (right).
Source: authors' calculations and construction

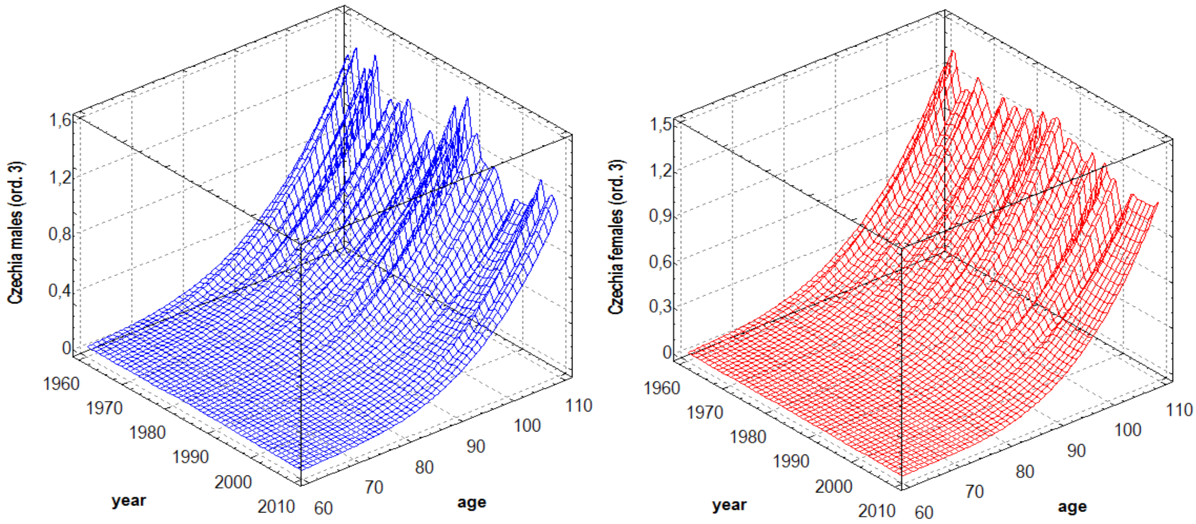

Fig. 5. Smoothing by polynomial function of the 3rd order for Czech males (left) and females (right).
Source: authors' calculations and construction

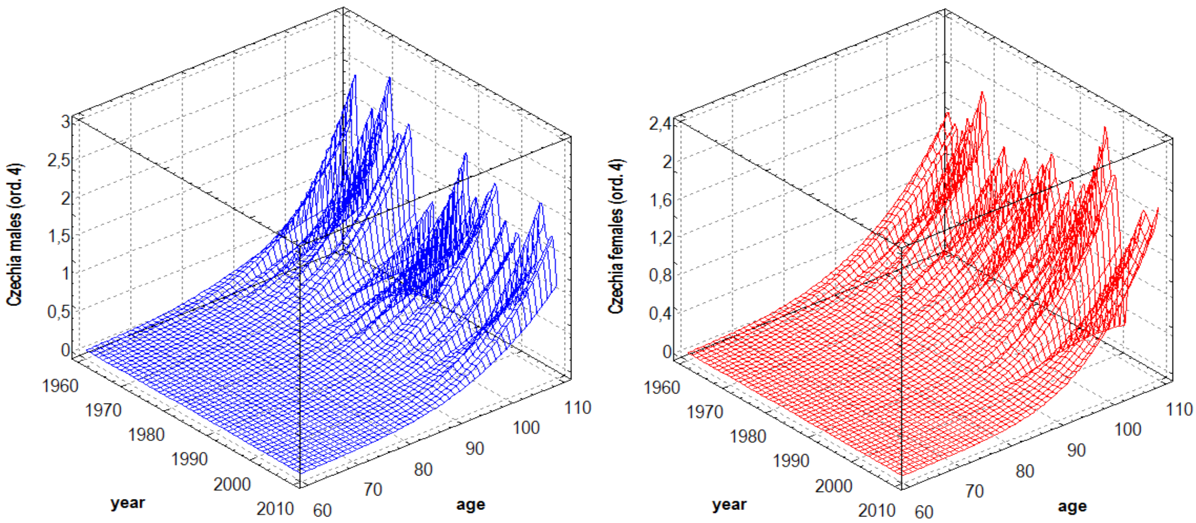

Fig. 6. Smoothing by polynomial function of the 4th order for Czech males (left) and females (right).
Source: authors' calculations and construction

A similar conclusion can be claimed in the case of the Czech population, where a polynomial function of 4th order works correctly in selected years only.

In the next part we dealt with the evaluation of selected models using the adjusted coefficient of determination. The evaluation was firstly carried out in Bulgarian population. Both in the case of males and females we can say, that for the modelling of intensity of mortality in the age range of 60–85 years is the most appropriate polynomial function of the 3rd or 4th order, but towards the present, as seems to be more appropriate polynomial function of the 3rd order. In the case of the Czech population we came to a similar conclusion. Here is the most appropriate model the polynomial function of the 3rd (or 4th, but not in general) order.

Acknowledgement

The authors gratefully acknowledge to the Czech Science Foundation project no. P402/12/G097 DYME – Dynamic Models in Economics for supporting this paper.

References

[1] BOLESLAWSKI, L., TABEAU, E. Comparing Theoretical Age Patterns of Mortality beyond the Age of 80. *In: Tabeau, E., Van Den Berg Jeths, A. and Heathcote, CH. (Eds.): Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective*, 2001, pp. 127 – 155.

[2] BURCIN, B., HULÍKOVÁ TESÁRKOVÁ, K., KOMÁNEK, D. *DeRaS: software tool for modelling mortality intensities and life table construction*. Charles University in Prague, 2012, Prague. URL: http://deras.natur.cuni.cz

[3] BURCIN, B., TESÁRKOVÁ, K., ŠÍDLO, L. Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci. *Demografie*, no. 52, vol. 2010, pp. 77 – 89.

[4] COALE, A., KISKER, E. Defects in Data on Old-age Mortality in the United States: New Procedures for Calculating Mortality Schedules and Life Tables at the Highest Ages. *Asian and Pacific Population Forum*, 1990, vol. 4, pp. 1 – 31.

[5] DOTLAČILOVÁ, P., ŠIMPACH, O., LANGHAMROVÁ, J. DeRaS versus MS Excel Solver in Levelling the Life Expectancy at Birth. *In: APLIMAT 2013*. Bratislava: Publishing House of STU, 2014, pp. 108 – 114.

[6] EKONOMOV, AL., YARIGIN, VN. The age dynamics of mortality and the Gompertz-Makeham law, *Zhurnal Obshchei Biologii*, 1989, vol. 50, no. 2, pp. 236 – 243.

[7] FIALA, T. Mortality in the Age from 20 to 70 Years in the Czech Republic in the Period 1950 to 1996. *Acta Oeconomica Pragensia*, 1999, vol. 7, no. 4, pp. 35 – 44.

[8] FIALA, T. Reducing the Consequences of Population Ageing for the Pension System by Raising the Retirement Age in the Czech Republic. *In: The 7th International Days of Statistics and Economics*. Slaný: Melandrium, 2013, pp. 398 – 407.

[9] GAVRILOV, L.A., GAVRILOVA, N.S. Mortality measurement at advanced ages: a study of social security administration death master file. *North American actuarial journal*, 2011, vol. 15, no. 3, pp. 432 – 447.

[10] GOMPERTZ, B. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London*, vol. 1825, no. 115, pp. 513 – 585.

[11] HMD: Wilmoth, J.R., Shkolnikov, V, Barbieri, M. *The Human Mortality Database*, URL: http://www.mortality.org/

[12] MAKEHAM, W.M. On the Law of Mortality and the Construction of Annuity Tables. *The Assurance Magazine, and Journal of the Institute of Actuaries*, vol. 1860, no. 8, pp. 301 – 310.

[13] ŠIMPACH, O., LANGHAMROVÁ, J. Czech Household Computer Facilities as a Reliable Variable in a Life Expectancy Forecast Model up to the Year 2060. *In: IDIMT-2012 (ICT Support for Complex Systems)*. Linz: Trauner, 2012, pp. 143 – 152.

[14] ŠIMPACH, O., LANGHAMROVÁ, J. The Impact of ICT Growth on Households and Municipalities in the Czech Nuts-3 Regions: The Application of Cluster Analysis. *In: IDIMT-2014 (Networking Societies – Cooperation and Conflict)*. Linz: Trauner, 2014, pp. 63 – 70.

[15] ŠIMPACH, O., PECHROVÁ, M. Development of the Czech Farmers' Age Structure and the Consequences for Subsidy Policy. *Agris on-line Papers in Economics and Informatics*, 2015, vol. VII, no. 3, pp. 57 – 69.

[16] THATCHER, R.A., KANISTÖ, V., VAUPEL, J.W. *The Force of Mortality at Ages 80 to 120*. Odense University Press, 1998.

Appendix

Following part presents the results of adjusted determination indices for particular populations.

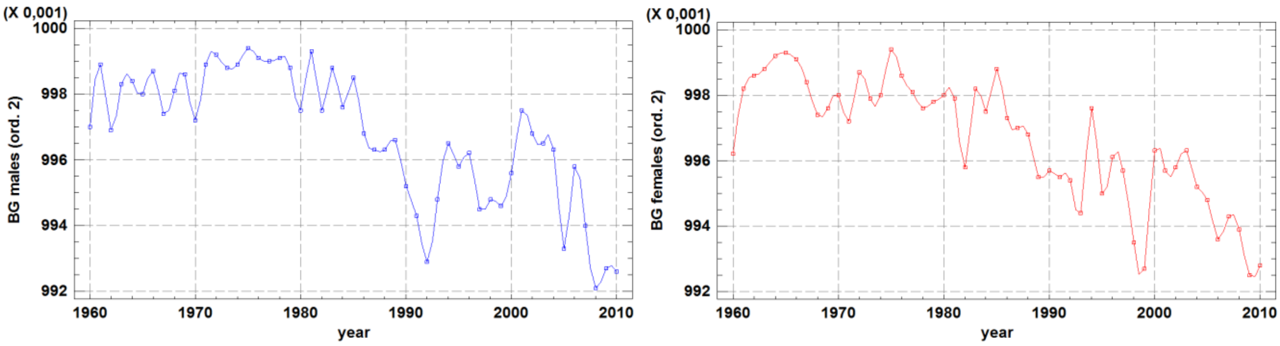

Fig. 7. Development of *adj.*R^2 for polynomial functions of the 2nd order for Bulgarian males (left) and females (right). Source: authors' calculations and construction

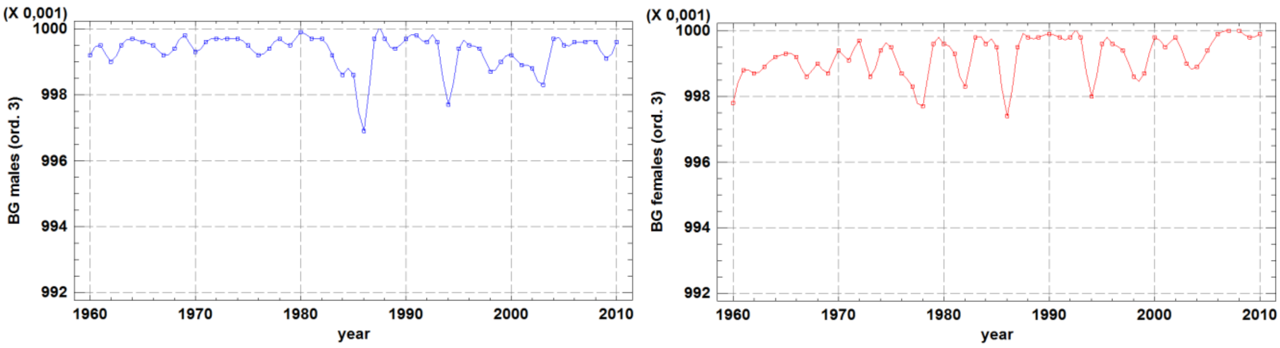

Fig. 8. Development of *adj.*R^2 for polynomial functions of the 3rd order for Bulgarian males (left) and females (right). Source: authors' calculations and construction

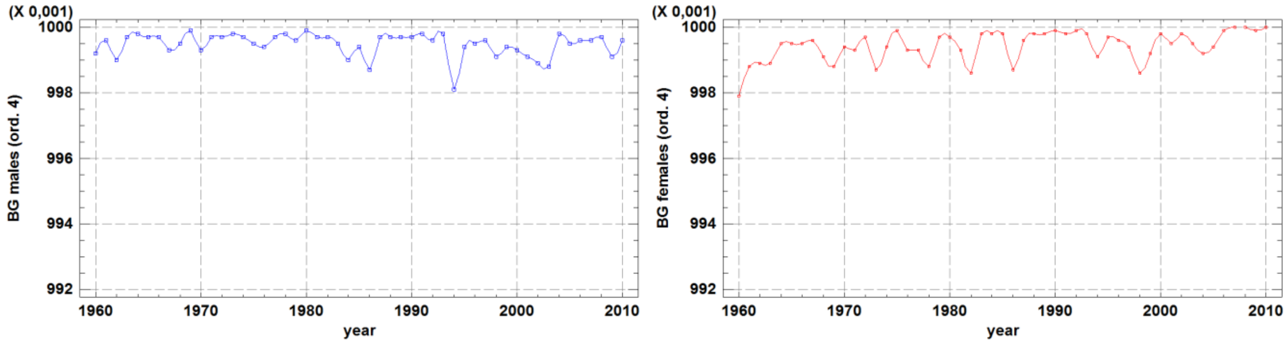

Fig. 9. Development of *adj.*R^2 for polynomial functions of the 4th order for Bulgarian males (left) and females (right). Source: authors' calculations and construction

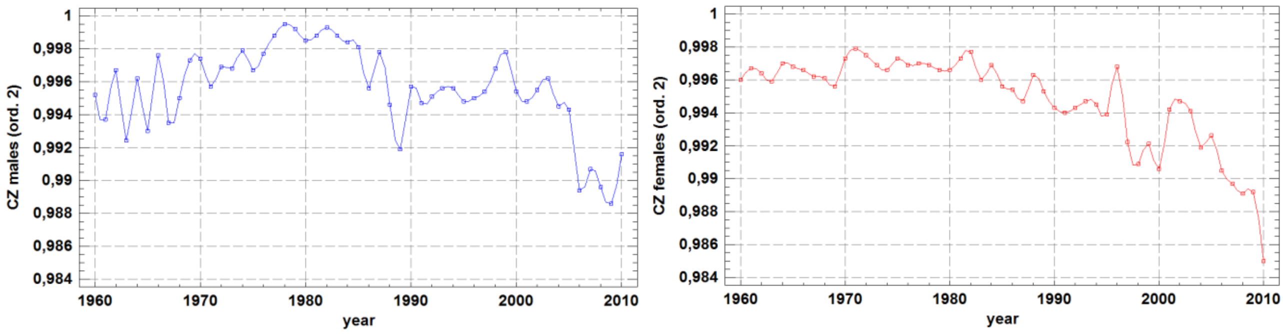

Fig. 10. Development of *adj.*R^2 for polynomial functions of the 2nd order for Czech males (left) and females (right). Source: authors' calculations and construction

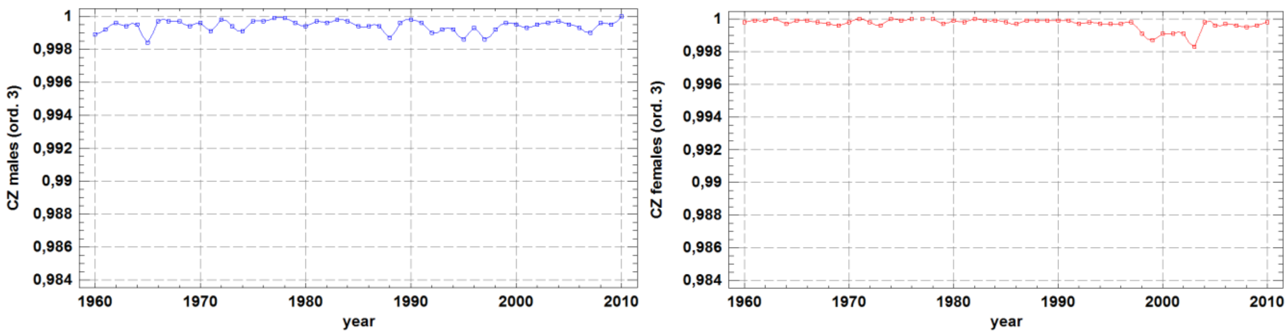

Fig. 11. Development of *adj.*R^2 for polynomial functions of the 3rd order for Czech males (left) and females (right). Source: authors' calculations and construction

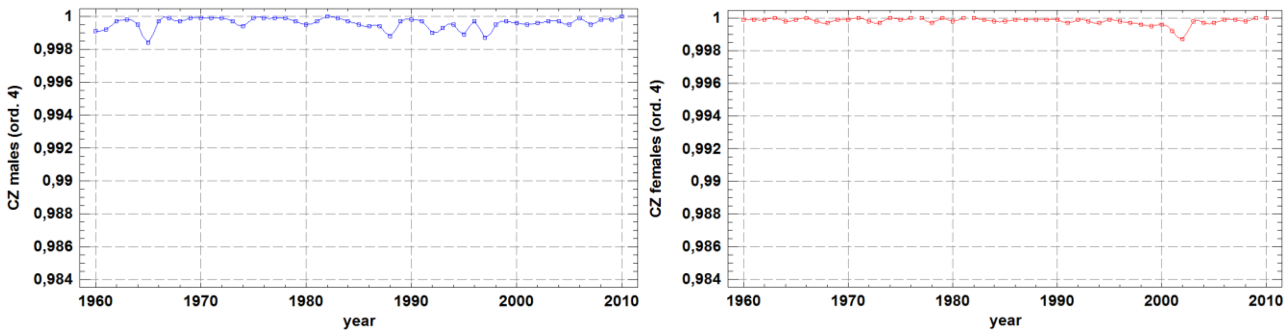

Fig. 12. Development of $adj.R^2$ for polynomial functions of the 4th order for Czech males (left) and females (right). Source: authors' calculations and construction

Current address

Petra Dotlačilová, Ing.
University of Economics in Prague, Faculty of Informatics and Statistics,
Department of Mathematics
W. Churchill sq. 4, 130 67 Prague 3, Czech Republic
Tel. number: +420 224 095 273, e-mail: petra.dotlacilova@vse.cz

Ondřej Šimpach, Ing., Ph.D.
University of Economics in Prague, Faculty of Informatics and Statistics,
Department of Statistics and Probability
W. Churchill sq. 4, 130 67 Prague 3, Czech Republic
Tel. number: +420 224 095 273, e-mail: ondrej.simpach@vse.cz