# Fitting and forecasting multi-population mortality models based on Hungarian regional data

**Lívia Varga**
Corvinus University of Budapest,
Hungarian Central Statistical Office,
Budapest, Hungary
E-mail: livia.varga@ksh.hu

The stochastic mortality model of Lee–Carter (1992) has become extremely popular over the past decades. As a result, many extended versions, for example, multi-population models, exist nowadays. In the original model, the logarithm of the mortality rates is a linear function of a time-dependent factor called the mortality index, which can be considered a long-term trend describing the improvement in mortality. The prediction of mortality rates is based on forecasting this long-term trend. Multi-population mortality models were developed to jointly model related populations.

This research fit five multi-population mortality models to Hungarian regional data by sex, and compared the models' results to each other and to the results of the original Lee–Carter model. This paper is the first to examine the possibility of applying multi-population mortality models to Hungarian regional data. The selected mortality models are discussed in a standardized methodological framework. The author use maximum likelihood estimation to observe the models on the base period of 1970–2021. Regarding the number of deaths, a Poisson distribution is assumed. As a result, the age-specific mortality rates are forecasted until 2050, and the future life expectancies of the regions of Hungary are estimated by sex. The author consider the presence of the COVID-19 pandemic during fitting and predicting the models similarly to Lee and Carter, who treated the effect of the Spanish flu in their study.

**Keywords:**
stochastic mortality models,
related populations,
Hungarian regions,
modeling jointly,
forecasting life expectancy

Regional Statistics, Vol. 13. No. 5. 2023: 863–898; DOI: 10.15196/RS130504

## Introduction

The study of Lee–Carter (1992) is a milestone in the probability theory of mortality. The most significant advantage of this model lies in its simplicity. The age-specific mortality rates can be predicted by considering only a few factors. Accounting for the number of deaths and population by age and calendar year is enough to fit these mortality models. The estimation does not require quantifying the past's medical, social, and other effects that impact mortality. Applying the mortality model of Lee and Carter, we assume that the long-term trend of mortality improvement continues in the future.

In the 30 years since the Lee–Carter study was conducted, the range of the stochastic mortality models has grown considerably wider. The multi-population mortality models related to the original Lee–Carter model may be appropriate for analyzing the mortality of subgroups (e.g., subpopulations with similar socioeconomic backgrounds). These models can be fitted, for example, to predict the mortality rates by territory. In this study, in addition to presenting the original model, we introduce and apply five modified versions to Hungarian regional data. These multi-population models incorporate common features of the subpopulations in addition to their group-specific parameters. Furthermore, we can quantify the mortality trend in multiple forms to differentiate the long-term effects of the past in the whole population and the short-term discrepancy from the main trend by groups.

Finding the most suitable regional forecasting model from the five fitted multi-population models requires comparing the different model variants and evaluating their applications on the Hungarian data. The results of the different multi-population models are also benchmarked to the original Lee–Carter model fitted by region. First, we review the theoretical background of the models and then introduce the data and the methodology of our analysis.

## Theoretical background

### The mortality model of Lee and Carter

The publication of **Lee–Carter** from 1992 is the most referenced study regarding stochastic mortality models. The authors analyzed the mortality data of the USA from 1900 to 1989 and prepared a forecast based on time series analysis. In their parsimonious model, the logarithm of central mortality rates is dependent on a time-specific mortality index, which is not observed. In general, this estimation does not require knowledge of past medical, social, and other effects that impact mortality. Instead, the model relies on a long-term, historical trend, which is the mortality index.

The initial equation is:

$$ln\, m_{xt} = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt} \quad (1)$$

where $m_{xt}$ is the central mortality rate, which equals $\frac{D_{xt}}{E_{xt}}$. This ratio denotes the number of deaths divided by the central population (exposure) by age $x$ and year $t$. The error term $\varepsilon_{xt}$ is white noise with zero mean and constant variance ($\sigma^2$), which refers to the effects not captured by the model. In the equation, the historical trend of the change (improvement) in mortality is described by $\kappa_t$. The parameter $\beta_x$ expresses the impact of $\kappa_t$ on mortality at different ages. Furthermore, $\alpha$ and $\beta$ depend on age, and $\kappa$ is dependent on time.

To achieve a unique solution during the estimation, Lee and Carter introduced constraints on $\beta_x$ and $\kappa_t$, which are the following: $\sum_x \beta_x = 1$ and $\sum_t \kappa_t = 0$. This also results in the $\alpha_x$ term being the average of $ln\, m_{xt}$ over time, so the $\alpha_x$ parameter refers to the general shape of the mortality. The $\beta_x$ coefficient could be negative for some ages (mortality increases at certain ages and decreases at other ages), but this will not be a problem in the long term. Negative mortality rates do not appear in the model, which is an advantage. The initial equation cannot be estimated with ordinary least squares (OLS) regression because there are an unknown index and other parameters on the right. Lee and Carter found the solution using the singular value decomposition (SVD) method.

The estimation of the future mortality rates can be provided by predicting $\kappa_t$. Thus, additional data from the life table (e.g., life expectancy) can be calculated as well. The projection of the mortality index indicates what would happen if the long-term trend continued. Analyzing the data of the USA, Lee and Carter found that the mortality index decreased approximately linearly between 1900 and 1989, so the mortality improved. They predicted the continuation of this trend until 2065. The authors found that $\kappa_t$ is a random walk with a drift process:

$$\kappa_t = \delta + \kappa_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\kappa^2) \quad (2)$$

where $\delta$ is the drift parameter, and $\epsilon_t$ is white noise.

## Some modifications of the Lee–Carter model

The Lee–Carter model has been criticized over the years, and a number of improved versions have been created. However, in the original study, Lee–Carter (1992) had already expanded the model. Due to the Spanish flu epidemic in 1918, a dummy was included in the equation for predicting the mortality index to filter out the impact of the flu. If we do not apply such a variable in the forecasting equation, we view the epidemic as an event that will reappear in the future. This affects only the confidence intervals and not the prediction itself. Without the dummy variable, the confidence interval would be wider. Due to the impact of the COVID-19 pandemic, it is advisable to use such a bivariate variable when we analyze the mortality data of recent years.

The study of Lee (2000) received some criticisms that had been made of the original model. For example, one of the weaknesses of the Lee–Carter estimation is

that past patterns do not necessarily prevail in the future. Structural changes may occur (e.g., in the field of medicine, but we can also mention the long-term presence and impact of a global epidemic). We can observe linearity in the time series of the mortality index in the 20th century. However, it is not certain whether this was the case in earlier periods of history or whether it will continue. A further problem may arise if the rate of decline in the mortality index differs across different age groups or if the assumption about the uncorrelated error terms between the cohorts is violated.

Today, there are several variants of the Lee–Carter model. Hunt–Blake (2014, 2015) and Villegas et al. (2018) introduce some important versions within a generalized framework. For example, it is possible to incorporate a cohort-specific factor as an extension of the model, but there are also models with more than one mortality index. The cohort effect refers to the fact that the change in mortality is not independent of the cohort.[1] Furthermore, the model estimation can be implemented separately by sex because of the differences between the mortality of men and women. We can also fit the mortality model by region or by cause of death.

## The Poisson Lee–Carter model

In the Lee–Carter model, the error term is homoscedastic and has a normal distribution. However, this is unrealistic because the number of deaths is smaller in the older age groups, and the logarithms of the mortality rates are more variable than in the younger ones. Brouhns et al. (2002) assumed the Poisson distribution of deaths and mentioned that, according to Brillinger (1986), this is an appropriate assumption regarding the number of deaths. We can fit the model of Brouhns et al. (2002) by applying the log function and central mortality rates. Similar to the original Lee–Carter model, a static age function ($\alpha_x$) and a nonparametric age–period factor ($\beta_x\kappa_t$) appear on the right side of the equation. Thus, the initial equation of the model is the same as Equation 1, but the number of deaths has a Poisson distribution:

$$D_{xt} \sim Poisson(E_{xt}m_{xt}). \tag{3}$$

The interpretation of the parameters and the two constraints related to the factor $\beta_x\kappa_t$ are the same as in the Lee–Carter model. Furthermore, Brouhns et al. (2002) assume that $\beta_1 = 1$. They use the maximum likelihood estimation method with Newton–Raphson iteration and provide the iterative updating scheme of the parameters $\alpha_x$, $\beta_x$, and $\kappa_t$. Instead of the ARIMA(0,1,0) model,[2] Brouhns et al. (2002) fitted ARIMA(0,1,1) to forecast the time-varying index:

$$\kappa_t = C + \kappa_{t-1} + \xi_t + \theta\xi_{t-1}, \qquad \xi_t \sim N(0, \sigma_\kappa^2) \tag{4}$$

where $C$ is a constant that expresses the average annual change in the mortality index leading to the long-term change in mortality, and $\xi_t$ is the independent disturbance.

[1] See Renshaw–Haberman (2006) who add the cohort effect to the Lee–Carter model.

[2] ARIMA is the short name of the autoregressive integrated moving average models (see, e.g., Kapás 2022, Khedhiri 2022).

## Multi-population mortality models

In this chapter, we focus on some multi-population mortality models, which belong to the extended family of the Lee–Carter model. In the long run, the independent forecast may unrealistically strengthen the divergence between the population groups in terms of age-specific mortality rates and life expectancy. If we would like to analyze population groups with similar socioeconomic backgrounds and mortality conditions, the choice of modeling jointly can be the most appropriate since it can help avoid divergence problems. Subgroups of a population may be, for example, different countries of a region, different territorial units of a country, or subgroups divided by sex. Given these subpopulations, it is reasonable to assume that mortality differences between the groups do not increase in the long run. Nondivergent forecasts for subgroups within a population are called coherent (Li–Lee 2005). In the following, we introduce the applied multi-population models of this study.[3] There are many other models in addition to these. For example, Villegas et al. (2017) provide an overview of the evolution of multi-population mortality models.

Analyzing the data of men and women, **Carter–Lee** (1992) proposed a common mortality index with an age-dependent coefficient that varies by sex. The initial equation of the model can be written as follows:

$$ln\, m_{xtj} = \alpha_{xj} + \beta_{xj}^{(1)}\kappa_t^{(1)} + \varepsilon_{xtj} \quad (5)$$

where $j$ refers to the subpopulation. The constraints are $\sum_x \beta_{xj}^{(1)} = 1$ for $\forall\, j$ and $\sum_t \kappa_t^{(1)} = 0$. However, a coherent forecast is not yet necessarily guaranteed by this variant. The group-specific $\beta_x$ can cause divergence problems.

Li–Lee (2005) introduce modifications of the Lee–Carter model to avoid divergence in the long term between the mortality of men and women. In this model, the mortality index and its coefficient are the same for each subpopulation. The initial equation is modified as follows:

$$ln\, m_{xtj} = \alpha_{xj} + \beta_x^{(1)}\kappa_t^{(1)} + \varepsilon_{xtj} \quad (6)$$

where the $\beta_x^{(1)}\kappa_t^{(1)}$ term is the common factor. The constraints are $\sum_x \beta_x^{(1)} = 1$ and $\sum_t \kappa_t^{(1)} = 0$. This version of the Lee–Carter model is called the common factor (**CF**) model by the authors. However, the first example of this idea is the work of Lee–Nault (1993). As mentioned by Li–Lee (2005), Lee–Nault, faced with the problem of divergence, had already fitted the common factor mortality model in their territorial model of Canada. The $\alpha_{xj}$ term in the model of Li–Lee (2005) is determined separately for each population. This does not cause divergence in the long run. The authors use the SVD method and assume a homoscedastic error term.

[3] The notations used for the model parameters may differ from those used by the authors. We apply superscript in the case of the factor $\beta_x\kappa_t$ to identify the different indices and their coefficients easily in the models with multiple factors. We denote the group-specific parameters by the subscript $j$.

Li–Lee argue for applying the CF model because the similarity of the subpopulations' socioeconomic background and the (expected permanent) close relationship between them justify the use of the common factor. Fitting the Lee–Carter model by subgroups of a population separately may be appropriate if the time series of the mortality indices have the same drift and the $\beta_x$ coefficients are identical. In this case, the ratios of the subpopulations' mortality rates will be constant over time for all ages in the projection. However, it is unlikely that two or more different mortality indices have the same drift. In practice, it is a sufficient condition that the groups of the population have common $\beta_x$ and $\kappa_t$.

An additional time-varying factor can be included in the model to account for the periodic deviation from the long-term trend, which may differ across groups. The common factor is estimated on the data of the whole population in the same way as in the original Lee–Carter method. This variant is the augmented common factor (**ACF**) model, which is also presented by Li–Lee (2005). The initial equation is:

$$ln\, m_{xtj} = \alpha_{xj} + \beta_x^{(1)}\kappa_t^{(1)} + \beta_{xj}^{(2)}\kappa_{tj}^{(2)} + \varepsilon_{xtj} \quad (7)$$

where the second factor $\beta_{xj}^{(2)}\kappa_{tj}^{(2)}$ allows the short- and medium-term discrepancy from the mortality change indicated by the common factor. If these differences were permanent in the long run, the forecast would be divergent. Thus, in this extended model, the common factor specifies the long-term trend for the whole population, and the second factor represents the shorter-run deviations, which are specific for each population $j$. First, Li and Lee fit the original Lee–Carter model to the aggregated data of the subgroups, and then in the next step, the additional factor is derived from the residual matrix[4] of the common factor model applying the SVD method. The restrictions of the parameters are $\sum_x \beta_x^{(1)} = 1$, $\sum_t \kappa_t^{(1)} = 0$, $\sum_x \beta_{xj}^{(2)} = 1$ for $\forall\, j$, and $\sum_t \kappa_{tj}^{(2)} = 0$ for $\forall\, j$.

According to Li–Lee (2005), we can assume the independence of the factors $\kappa_t^{(1)}$ and $\kappa_{tj}^{(2)}$ and the independence of the second factor between the groups because they refer to random changes in different populations. For both common and group-specific factors, the random walk process can be used in the projection, and the authors also mention the possibility of fitting the AR(1) model for the second index.

Li (2013) modified the extended CF model of Li and Lee. On the one hand, the innovation is to assume a Poisson distribution for the number of deaths and estimate with the maximum likelihood method following the approach of Brouhns et al. (2002). On the other hand, Li (2013) incorporates more than one additional group-specific factor into the model. The author analyzed the Australian data by two subpopulations (men and women). Li (2013) argues for the ACF model. If the trend differences observed in the past among the groups are expected to continue (because of, e.g., biological endowments), then a coherent model is needed. For example, as

[4] $ln\, m_{xtj} - \alpha_{xj} - \beta_x^{(1)}\kappa_t^{(1)}$

mentioned by Li (2013), infant mortality rates are usually lower for girls than for boys in developed countries, and it is implausible to imagine the opposite occurring in the future. In projections for subpopulations, some trends can be assumed to continue. The ACF model may also guarantee avoiding the crossover effect. Assuming $n$ additional factors and Gaussian error structure, the main equation of the model is:

$$ln\, m_{xtj} = \alpha_{xj} + \beta_x^{(1)}\kappa_t^{(1)} + \sum_{i=2}^{n} \beta_{xj}^{(i)}\kappa_{tj}^{(i)} + \varepsilon_{xtj}.^{5} \quad (8)$$

However, applying the model with too many parameters should be avoided, according to Li (2013). The author distinguishes at most six additional factors. However, if the third- or higher-order indices appear as irregular components, we should decrease the number of factors that are less suitable for prediction. The author suggests AR(p) models for the forecast of the additional time-varying factors. Furthermore, he uses the random walk with drift model for the common mortality index. The error terms of the time series are assumed to be independent from each other and across time. In the case of the ACF model, the following constraints are applied: $\sum_x \beta_x^{(1)} = 1$, $\sum_t \kappa_t^{(1)} = 0$, $\sum_x \beta_{xj}^{(i)} = 1$ for $\forall\, i,\ j$, and $\sum_t \kappa_{tj}^{(i)} = 0$ for $\forall\, i,\ j$. According to Li (2013), model fit can be improved by incorporating multiple factors into the equation.

Li et al. (2016) mentioned the possibility of using common age-varying coefficients for the subpopulations following the common age effect (**CAE**) model of Kleinow (2015). This variant is called the augmented common factor model with a common age effect (**ACF–CAE**). The estimated equation is:

$$ln\, m_{xtj} = \alpha_{xj} + \beta_x^{(1)}\kappa_t^{(1)} + \sum_{i=2}^{n} \beta_x^{(i)}\kappa_{tj}^{(i)} + \varepsilon_{xtj}. \quad (9)$$

Related to the initial equation, the constraints are $\sum_x \beta_x^{(1)} = 1$, $\sum_t \kappa_t^{(1)} = 0$, $\sum_x \beta_x^{(i)} = 1$ for $\forall\ i$, and $\sum_t \kappa_{tj}^{(i)} = 0$ for $\forall\, i,\ j$. If $\beta_{xj}^{(i)}$ is very similar in the subpopulations, then it is particularly important to implement the common age-varying coefficients. As described by Li et al. (2016), we can determine the optimal number of additional factors based on the Bayesian Information Criterion (BIC) values, the residual plots, the trends of the additional indices, and the amount of available data. We can avoid overparameterization by analyzing the results from these aspects. According to Li et al. (2016), the vector autoregressive (VAR) model can be an alternative method to forecast the additional time-varying indices by capturing the relationships between them.

Kleinow (2015) attempted to use the common age effect in the context of the Lee–Carter mortality model without additional factors and explained this by the similar socioeconomic background of the subpopulations, which may result in a very similar age structure. The author analyzed developed countries with different

[5] The error term is nonadditive in the case of a Poisson distribution of deaths.

population sizes. Kleinow's modification leads to the following change in the Lee–Carter model:

$$ln\, m_{xtj} = \alpha_{xj} + \beta_x^{(1)}\kappa_{tj}^{(1)} + \varepsilon_{xtj} \quad (10)$$

where the mortality index differs between the groups, but the age effect is identical. The constraints are the following: $\sum_x \beta_x^{(1)} = 1$ and $\sum_t \kappa_{tj}^{(1)} = 0$ for $\forall\, j$. Kleinow applies the central mortality rates and the common principal component analysis method. First, he estimates the common factor ($\beta_x^{(1)}\kappa_t^{(1)}$) using the SVD. As mentioned by Kleinow (2015), we can create groups based on clustering the age effects of the subpopulations. In this case, each group would have common $\beta_x$, but these age effects would be different between the groups. This variant can be a further extension of the CAE model. Depending on which estimated parameters are considered common for the subpopulations, it is possible to use additional model versions regarding the simple or the augmented form (see, e.g., Enchev et al. 2017, Wen et al. 2021). In the following chapters, we fit and forecast five multi-population mortality models based on Hungarian regional data.[6] We examine how these models can be applied to predict life expectancy at birth for men and women by region. We compare the results of all multi-population models with each other and with those of the Lee–Carter model. First, we present the corresponding data and methodology. Then, we introduce the fitted parameters of the models and check the goodness of fit. After that, we will prepare the forecast of the mortality indices and calculate future life expectancy at birth by sex and region.

## Data and methodology

In this study, we analyze the mortality and population data of the Hungarian Central Statistical Office (HCSO). Annual age-specific records by region are available in this database from 1970 onward. While the longest possible time series would be preferred in this analysis, regions have not yet digitalized data for earlier years. Consequently, we analyze age-specific data, but the cohort of 90-year-olds is an aggregated group. The number of population aged above 90 is only available since 2012 by age in years, so we have no detailed information about this cohort in the whole base period. Throughout the study, 2021 marks the last observed year. As mortality rates for men and women differ largely, we fit the mortality models separately by sex.

Adjusted population numbers are used in this analysis for the 1980s and 1990s. Starting from the latest census, the size of the population is determined by considering the monthly vital events and the number of net migrations. However, the population number calculated in this way is generally different from the result of the next population census (in the year of the census) due to the uncertainty in the data.

[6] Danesi et al. (2015) and Scognamiglio (2022) also compare several multi-population extensions of the Lee–Carter model on Italian regional data.

Therefore, a retrospective correction is made based on the latest census. For the 1980s and 1990s, the counties' total population number adjustment was also calculated in the years 1990 and 2001 after the two censuses (see HCSO 1990, 2001). However, the recalculation has not been made by sex and age. For this reason, we estimated the distribution by sex and age in the case of the recalculated data using the population numbers, which rely on the monthly statistics of the vital events. The data for deaths of unknown age and/or region are not ignored in this study. Instead, we examined the distribution of these data proportionately considering the detailed information of known data. After this correction, the number of age-specific death counts was rounded to the nearest integer.

## Descriptive statistics

Analysis of the mortality data confirms the differences between men and women. Figure 1 illustrates the logarithm of the smoothed mortality rates by sex in Hungary over past decades. For both sexes, it can be seen that overall mortality improved because the curve gradually shifted down. The most negligible improvement is observed among the 55–65-year-old people, while the younger age groups show the most significant improvement in mortality. It can also be concluded that log mortality rates are lower for women than men. Generally, the mortality of females is more favorable, especially in the cohort of 20-year-olds. In particular, traffic accidents can play a role in this difference between men and women (for example, a newly obtained driver's license is a greater threat to males than females).[7] Furthermore, it is important to highlight that in the first half of the 1990s, political and economic changes had a negative impact on mortality, especially for middle-aged people.

Figure 1 does not show the data for 2020 and 2021. However, these years are crucial because of the spread of the COVID-19 epidemic (Kincses–Tóth 2020). Figures 2–3 already visualize the differences in mortality by region for these years. Figure 2 presents the evolution of the age-standardized mortality rates (ASMRs) for men and women. This aggregated measure is calculated by comparing the number of deaths ($D_{xtj}$) to the mid-year population ($E_{xtj}$) multiplied by the age-varying weights of the standard population. According to Wen et al. (2021), the formula of the ASMR can be expressed as follows:

$$ASMR_{tj} = \sum_{x\in\chi} \frac{D_{xtj}}{E_{xtj}} w_x, \quad w_x = \frac{E_x^S}{\sum_{x\in\chi} E_x^S} \tag{11}$$

where the weights $w_x$ refer to the age-specific distribution of the standard population $E_x^S$. In this study, we chose central exposure in 1970 as a reference for the standard population. We examined the values of the ASMR by region $j$ and sex separately. The age range $\chi$ denotes the age in years, which is between 0 and 90.

[7] See, e.g., the Tables 6.2.20 and 6.2.21 in HCSO (2021).

Figure 1

The logarithm of the central mortality rates in Hungary

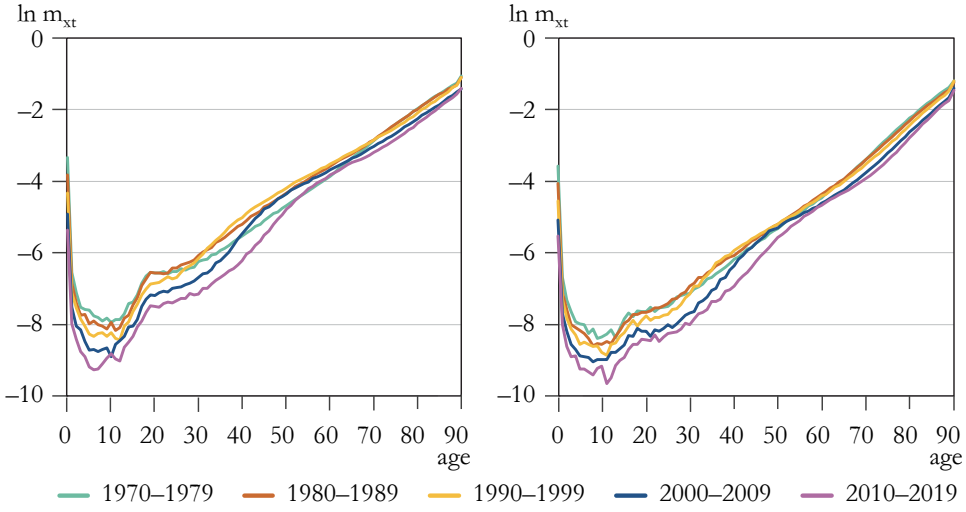



Figure 2

Age-standardized mortality rates by region

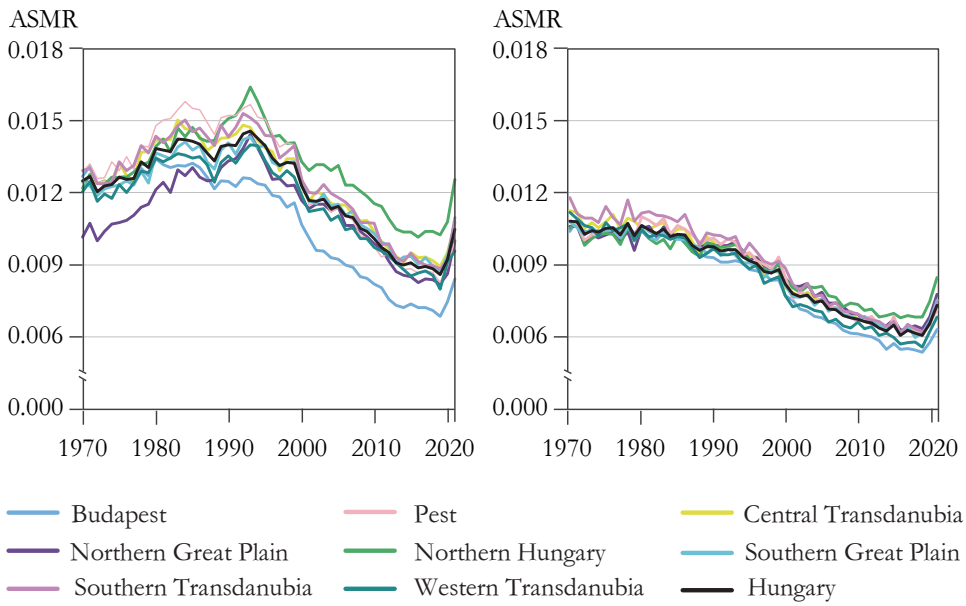
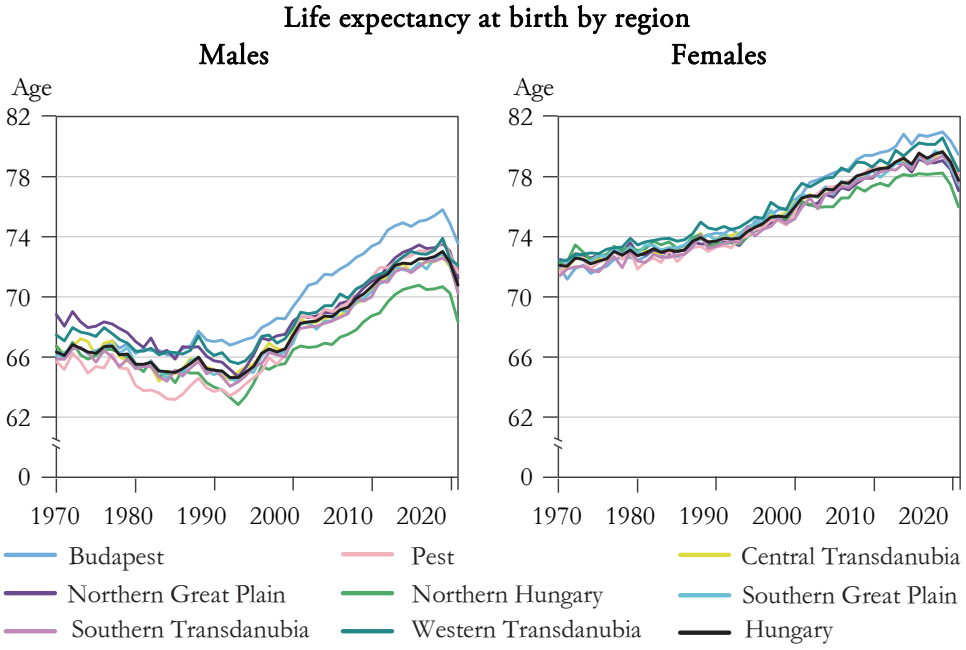


Figure 3




Based on the time series values of the ASMR, it is noticeable that the higher mortality in the 1990s had a greater impact on men than women, especially in Northern Hungary and the Northern Great Plain. The primary trend is that males were characterized by declining mortality from the 1970s to the mid-1990s. In their case, long-term improvement began only after these decades, though this improvement ceased upon the arrival of the COVID-19 epidemic. For women, the curve of the ASMR values stagnated from 1970 and decreased after the mid-1980s regarding the long-term trend. The ASMR values by region are more similar for women, and the differences between the regions are more marked in the case of men.[8] COVID-19 has influenced the mortality of both sexes. These findings are also supported by the time series of life expectancy at birth[9] (see Figure 3). The evolution of these curves is contrary to the observed ASMR values, as the higher mortality rate (the high number of deaths in a given year) results in a lower life expectancy. Figure 3 shows more clearly that gradual improvement in mortality can be seen for females from the 1970s. The national level of life expectancy at birth was 73.1 years for men and 79.7 years for women in 2019 before the COVID-19 pandemic. These values

[8] Analyzing life expectancy, Bálint (2011) has also found that territorial differences are more pronounced for men than for women.

[9] We examined the life table made by Chiang (1968), but we chose 0.2 as the average fraction of the year lived by an infant, and the fraction of the last year of life was 0.5 for all other cohorts following the suggestion of Eurostat (see https://ec.europa.eu/eurostat/cache/metadata/Annexes/demo_mor_esms_an1.pdf).

have been reduced to 70.8 and 77.8 years in 2021. Life expectancy was the highest in Budapest and the lowest in Northern Hungary for both sexes in recent years.

The main trends of ASMR and life expectancy are in line with some important findings of Józan (2008, 2009, 2012) and Bálint (2016). Three periods of epidemiological development are distinguished from the second half of the 20th century, according to Józan (2008). The second and third of these are in the scope of this study. The period between 1948 and 1966 (the 1st phase) was the hopeful beginning. Infant and child mortality was substantially reduced, and the life expectancy of young adults was also extended. The development of successful cures for pneumonia, tuberculosis, and other infectious diseases was the main contributor to an increase in life expectancy.

The years 1967–1993 (the 2nd phase) were characterized by a chronic, qualified epidemiological crisis. An epidemiological crisis means that mortality increased and that life expectancy at birth decreased. Being chronic refers to that the condition existed constantly for decades. Being qualified means that this crisis hit only a subgroup (middle-aged men) of the population. Meanwhile, the mortality improved among women because the losses of middle-aged women's life expectancy were offset by the gains of young and old cohorts. Bálint (2016) In the 2nd phase, life expectancy decreased due to neoplasms, cardiovascular and digestive diseases in addition to violent deaths. Lethal diseases and the lifestyle that causes them were already typical in the 1950s, and they left their mark on the mortality of the 1960s and 1970s. The epidemiological crisis peaked in 1993 (Józan 2012).

The political and economic change of the regime was needed to reverse this trend in mortality. The 3rd phase, beginning in 1994, was the period of renewal. During this phase, the improvement in infant and child mortality slowed, and life expectancy for the middle-aged and the elderly populations increased. High blood pressure, cerebrovascular and ischemic heart diseases, and diabetes mellitus were effectively treated (Józan 2008). Chronic noninfectious diseases were dominant in the 3rd phase, but these diseases appeared later in life, and the progression of diseases improved (Józan 2009). The spread of health awareness, effective medicine, medical technology, and emergency care played a significant role in decreasing mortality (Józan 2012).

## The fitted mortality models

If we would like to analyze mortality without considering national characteristics (trend and/or age effects), we can rely on the results of the Lee–Carter model. Table 1 summarizes all the mortality models included in this study. In the case of the ACF and the ACF–CAE models, we calculate with one additional factor. In total, six different mortality models are fitted on the base period between 1970 and 2021. In each case, we assume the Poisson distribution of deaths and consider central mortality rates. The models are fitted by applying maximum likelihood estimation with the

Newton–Raphson iteration method to determine the parameter values. The iteration continues until the change in the log-likelihood value is less than $10^{-6}$.

We use the R software (R Core Team 2021) for the numerical programming of the models and other computations. In the case of the multi-population mortality models, we first examine the parameters for the whole population. Then, we incorporate these parameters as known values into the models of the subpopulations. After this, the group-specific parameters are estimated. The algorithm is shown here. All models are estimated by age in years, but the mortality data for the age group of 90 and above are aggregated. In addition to the previously mentioned reason, this is also practical because the uncertainty of the estimation may be increased for smaller groups. Mortality rates are generally more volatile among the population aged above 90.

Table 1

**The fitted mortality models based on Hungarian regional data**

| Literature | The names of the models | The initial equations |
|---|---|---|
| Lee–Carter (1992), Brouhns et al. (2002) | Lee–Carter | $ln\, m_{xtj} = \alpha_{xj} + \beta_{xj}^{(1)}\kappa_{tj}^{(1)}$ |
| Carter–Lee (1992) | Carter–Lee | $ln\, m_{xtj} = \alpha_{xj} + \beta_{xj}^{(1)}\kappa_{t}^{(1)}$ |
| Lee–Nault (1993), Li–Lee (2005) | CF | $ln\, m_{xtj} = \alpha_{xj} + \beta_{x}^{(1)}\kappa_{t}^{(1)}$ |
| Kleinow (2015) | CAE | $ln\, m_{xtj} = \alpha_{xj} + \beta_{x}^{(1)}\kappa_{tj}^{(1)}$ |
| Li–Lee (2005), Li (2013) | ACF | $ln\, m_{xtj} = \alpha_{xj} + \beta_{x}^{(1)}\kappa_{t}^{(1)} + \beta_{xj}^{(2)}\kappa_{tj}^{(2)}$ |
| Li et al. (2016) | ACF–CAE | $ln\, m_{xtj} = \alpha_{xj} + \beta_{x}^{(1)}\kappa_{t}^{(1)} + \beta_{x}^{(2)}\kappa_{tj}^{(2)}$ |

*Note:* The subscript $j$ refers to the subpopulation.

## Results

In the first step, the different mortality models were fitted to adjusted and nonadjusted population data, and then we analyzed the differences between the parameters. We found that the age-specific, regional population correction dampened the outlier values of the mortality indices between 1980 and 2000. In this study, we show the results of the fitted models based on the adjusted population. We have selected one of the eight regions to introduce the most important characteristics and statistics of the mortality models. Looking at the long time series of life expectancy at birth, the Southern Great Plain accurately represents the national mortality levels for both men and women. Therefore, we have chosen this territory as a representative region. For more information about the rest of the regions, figures and statistics are available here.

There were no convergence problems during the iteration. However, we did not accept the parameter estimation results regarding the Carter–Lee model for Northern Hungarian women.

## The fitted parameters

Figure A1 in the Appendix presents the values of the mortality models' fitted parameters for men and women separately in the Southern Great Plain. In the case of the models that include common trend and/or age effect(s) in addition to the subpopulation characteristics, the parameters derived from the national level estimation are equal. The values of $\alpha_{xj}$ are different for each model because they are slightly modified during the estimation procedure due to the fixed, national-level parameters. We chose the mean of the log mortality rates between 1970 and 2021 by age as initial values for the term $\alpha_{xj}$.

The year-to-year trend in mortality (the change in $\kappa_{tj}^{(1)}$ and $\kappa_{tj}^{(2)}$) has been volatile over the past decades. Overall, the long-term improvement in mortality (see the figure of the parameter $\kappa_{tj}^{(1)}$) for men started in the second half of the 1990s, although this has reversed in recent years. According to the results of the ACF and the ACF–CAE models, we can conclude similar findings for women. In the case of the Lee–Carter, Carter–Lee, and CF models, a continuously improving trend for females can be observed until 2019. The decline in 2020 and 2021 is the consequence of the COVID-19 epidemic's effect on the number of deaths for both men and women. For the ACF and the ACF–CAE models, the values of the additional mortality index $\kappa_{tj}^{(2)}$ are not irregular, so they appear predictable for both sexes. In the models with two factors, the shapes of the curves of the two mortality indices differ, confirming that it may be worthwhile to quantify the group-specific factor in addition to the national trend.

The values of the terms $\beta$ are not only positive regarding the different mortality models and subpopulations. Negative values appear in the case of the Lee–Carter model for women in Pest and for men in Northern Hungary. According to the Carter–Lee model, there is also one negative value for 9-year-old females in Pest. Negative values indicate increasing mortality for the relevant ages and regions. Furthermore, some negative values are fitted for the group-specific parameters $\beta_{xj}^{(2)}$ of the ACF model related to both men and women in certain regions. Since this model has two factors and there are no negative values for the common parameter $\beta_x^{(1)}$, mortality will not necessarily increase in the future. Suppose there were additional constraints in the models for the parameters $\beta$ to exclude the appearance of the negative values. In that case, the increasing mortality could be avoided for all ages provided that the predicted values of the parameters $\kappa$ are not positive. However, such constraints are not applied in this analysis.

## The goodness of fit

There are several measures related to the goodness of fit, which can support the choice between the models. For example, we can compare the values of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The lower the AIC or the BIC value is, the better the model is. The AIC and BIC values can be calculated as follows (see, e.g., Villegas et al. 2018):

$$AIC = 2k - 2L, \tag{12}$$

$$BIC = k\,ln(n) - 2L \tag{13}$$

where $k$ is the number of parameters, $n$ refers to the number of observations, and $L$ is the maximum log-likelihood value. The mean absolute percentage error (MAPE) is also a widely used measure to evaluate the accuracy of the different mortality models (see, e.g., Li 2013). We prefer the lowest MAPE value during the comparison of the models.

$$MAPE_j = \frac{1}{n}\sum\nolimits_{xt}\left|\frac{\hat{m}_{xtj} - D_{xtj}/E_{xtj}}{D_{xtj}/E_{xtj}}\right| \tag{14}$$

where $n$ denotes the number of observations, $D_{xtj}$ is the observed number of deaths, $E_{xtj}$ refers to the central exposure, and $\hat{m}_{xtj}$ is the model-specific, fitted mortality rate at age $x$, year $t$, and subpopulation $j$. The explanation ratio (see, e.g., Wen et al. 2021, Li–Lee 2005) is also a possible measure of the goodness of fit, which can be determined in the following way:

$$R_j = 1 - \frac{\sum_{xt}(ln\frac{D_{xtj}}{E_{xtj}} - ln\,\hat{m}_{xtj})^2}{\sum_{xt}(ln\frac{D_{xtj}}{E_{xtj}} - \alpha_{xj})^2} \tag{15}$$

where $D_{xtj}$, $E_{xtj}$, and $\hat{m}_{xtj}$ denote the same as before. The term $\alpha_{xj}$ in the denominator refers to the average of the log death rates over the years of the base period by age.

$$\alpha_{xj} = \frac{1}{n_Y}\sum_{t} ln\frac{D_{xtj}}{E_{xtj}} \tag{16}$$

where $n_Y$ is the total number of calendar years. In this study, we apply central exposures to calculate the explanation ratio. The higher the value for a mortality model is, the better the goodness of fit is.

Tables 2 and 3 show the values of the AIC, BIC, and MAPE related to the fitted mortality models for all regions. For males, the ACF model is the most appropriate, and the ACF–CAE model is the second best choice. We can declare similar conclusions for females based on most of the measures. Overall, the additional factor reduces the AIC, BIC, and MAPE values and increases the explanation ratio, so the second time-varying index is justified. The outliers are the BIC values for women, which are the lowest for the Lee–Carter model (and not for the augmented models) in all regions except Budapest.

Table 2

## The mortality models' goodness of fit for males, 1970–2021

| Regions | Measures | Lee–Carter | Carter–Lee | CF | CAE | ACF | ACF–CAE |
|---|---|---|---|---|---|---|---|
| Budapest | AIC | 5,318,727 | 5,320,350 | 5,322,359 | 5,319,795 | **5,313,249** | *5,314,690* |
| | BIC | 5,320,226 | 5,321,849 | 5,323,859 | 5,321,294 | **5,315,660** | *5,317,101* |
| | MAPE | 0.225 | 0.227 | 0.237 | 0.228 | **0.191** | *0.197* |
| | R | *0.499* | 0.512 | 0.481 | 0.427 | **0.580** | 0.487 |
| Central Transdanubia | AIC | 3,172,025 | 3,172,114 | 3,172,185 | 3,172,103 | **3,168,282** | *3,168,355* |
| | BIC | 3,173,524 | 3,173,614 | 3,173,685 | 3,173,602 | **3,170,692** | *3,170,765* |
| | MAPE | 0.229 | 0.231 | 0.232 | 0.231 | **0.193** | *0.195* |
| | R | 0.382 | 0.376 | 0.372 | 0.378 | **0.486** | *0.483* |
| Northern Great Plain | AIC | 4,150,538 | 4,152,183 | 4,153,092 | 4,152,339 | **4,146,637** | *4,147,339* |
| | BIC | 4,152,037 | 4,153,682 | 4,154,592 | 4,153,839 | **4,149,047** | *4,149,750* |
| | MAPE | 0.240 | 0.230 | 0.236 | 0.237 | **0.191** | *0.203* |
| | R | 0.344 | 0.394 | 0.374 | 0.376 | **0.521** | *0.492* |
| Northern Hungary | AIC | 3,998,530 | 4,000,420 | 4,000,786 | 3,999,456 | **3,992,801** | *3,993,114* |
| | BIC | 4,000,030 | 4,001,920 | 4,002,285 | 4,000,955 | **3,995,212** | *3,995,524* |
| | MAPE | 0.251 | 0.238 | 0.241 | 0.237 | **0.187** | *0.193* |
| | R | 0.242 | 0.221 | 0.262 | 0.310 | **0.477** | *0.464* |
| Pest | AIC | 3,022,995 | 3,024,011 | 3,024,489 | 3,023,642 | **3,019,225** | *3,019,943* |
| | BIC | 3,024,495 | 3,025,510 | 3,025,988 | 3,025,141 | **3,021,635** | *3,022,354* |
| | MAPE | 0.240 | 0.254 | 0.256 | 0.246 | **0.208** | *0.214* |
| | R | 0.518 | 0.513 | 0.507 | 0.513 | **0.587** | *0.582* |
| Southern Great Plain | AIC | 4,384,221 | 4,384,292 | 4,384,632 | 4,384,487 | **4,379,021** | *4,379,297* |
| | BIC | 4,385,720 | 4,385,791 | 4,386,131 | 4,385,986 | **4,381,431** | *4,381,708* |
| | MAPE | 0.218 | 0.218 | 0.222 | 0.223 | **0.180** | *0.186* |
| | R | 0.443 | 0.440 | 0.427 | 0.433 | **0.548** | *0.532* |
| Southern Transdanubia | AIC | 3,068,996 | 3,069,075 | 3,069,282 | 3,069,208 | **3,064,788** | *3,064,942* |
| | BIC | 3,070,495 | 3,070,574 | 3,070,781 | 3,070,707 | **3,067,199** | *3,067,352* |
| | MAPE | 0.236 | 0.236 | 0.237 | 0.237 | **0.198** | *0.199* |
| | R | 0.350 | 0.344 | 0.340 | 0.335 | **0.461** | *0.452* |
| Western Transdanubia | AIC | 2,954,563 | 2,954,651 | 2,954,773 | 2,954,694 | **2,951,070** | *2,951,213* |
| | BIC | 2,956,062 | 2,956,150 | 2,956,272 | 2,956,193 | **2,953,480** | *2,953,624* |
| | MAPE | 0.243 | 0.244 | 0.245 | 0.244 | **0.205** | *0.206* |
| | R | 0.336 | 0.333 | 0.327 | 0.329 | **0.441** | *0.433* |

*Note:* The most favorable values are highlighted in bold, and the second-best choice is in italics.

Table 3

## The mortality models' goodness of fit for females, 1970–2021

| Regions | Measures | Lee–Carter | Carter–Lee | CF | CAE | ACF | ACF–CAE |
|---|---|---|---|---|---|---|---|
| Budapest | AIC | 5,815,269 | 5,816,891 | 5,816,621 | 5,815,625 | **5,814,146** | *5,814,635* |
| | BIC | *5,816,768* | 5,818,390 | 5,818,120 | 5,817,124 | **5,816,556** | 5,817,045 |
| | MAPE | 0.222 | 0.238 | 0.242 | 0.231 | **0.207** | *0.219* |
| | R | 0.464 | 0.467 | 0.444 | 0.447 | **0.504** | *0.479* |
| Central Transdanubia | AIC | 2,706,458 | 2,706,541 | 2,706,661 | 2,706,583 | **2,706,165** | *2,706,293* |
| | BIC | **2,707,957** | *2,708,040* | 2,708,160 | 2,708,082 | 2,7085,75 | 2,708,703 |
| | MAPE | 0.236 | 0.236 | 0.236 | 0.237 | **0.222** | *0.223* |
| | R | 0.250 | 0.258 | 0.265 | 0.262 | **0.304** | **0.304** |
| Northern Great Plain | AIC | 3,903,348 | 3,903,488 | 3,903,683 | 3,903,426 | **3,902,511** | *3,902,650* |
| | BIC | **3,904,847** | 3,904,987 | 3,905,182 | 3,904,925 | *3,904,921* | 3,905,060 |
| | MAPE | 0.239 | 0.240 | 0.241 | 0.240 | **0.221** | *0.224* |
| | R | 0.339 | 0.330 | 0.340 | 0.349 | *0.388* | **0.398** |
| Northern Hungary | AIC | 3,523,379 | – | 3,524,586 | 3,523,570 | **3,522,820** | *3,523,079* |
| | BIC | **3,524,879** | – | 3,526,085 | *3,525,070* | 3,525,231 | 3,525,489 |
| | MAPE | 0.248 | – | 0.255 | 0.249 | **0.230** | *0.234* |
| | R | 0.195 | – | 0.127 | 0.193 | **0.256** | *0.240* |
| Pest | AIC | 2,640,869 | 2,641,056 | 2,641,247 | 2,641,158 | **2,640,338** | *2,640,558* |
| | BIC | **2,642,369** | *2,642,555* | 2,642,746 | 2,642,657 | 2,642,748 | 2,642,969 |
| | MAPE | 0.249 | 0.252 | 0.256 | 0.256 | **0.236** | *0.240* |
| | R | 0.343 | 0.345 | 0.325 | 0.327 | **0.378** | *0.371* |
| Southern Great Plain | AIC | 3,913,262 | 3,913,421 | 3,913,450 | 3,913,363 | **3,912,915** | *3,913,010* |
| | BIC | **3,914,762** | 3,914,920 | 3,914,949 | *3,914,862* | 3,915,325 | 3,915,421 |
| | MAPE | 0.235 | 0.236 | 0.237 | 0.237 | **0.222** | *0.225* |
| | R | 0.273 | 0.253 | 0.248 | 0.258 | **0.314** | *0.296* |
| Southern Transdanubia | AIC | 2,791,125 | 2,791,205 | 2,791,623 | 2,791,404 | **2,790,788** | *2,790,989* |
| | BIC | **2,792,624** | *2,792,704* | 2,793,122 | 2,792,903 | 2,793,198 | 2,793,399 |
| | MAPE | 0.240 | 0.241 | 0.241 | 0.240 | **0.227** | *0.229* |
| | R | 0.220 | 0.217 | 0.222 | 0.193 | **0.259** | *0.232* |
| Western Transdanubia | AIC | *2,630,525* | 2,630,669 | 2,630,891 | 2,630,653 | **2,630,405** | 2,630,550 |
| | BIC | **2,632,024** | 2,632,169 | 2,632,390 | *2,632,152* | 2,632,816 | 2,632,960 |
| | MAPE | 0.234 | 0.235 | 0.236 | 0.235 | **0.222** | *0.224* |
| | R | 0.259 | 0.276 | 0.262 | 0.232 | **0.299** | *0.267* |

*Note:* The most favorable values are highlighted in bold, and the second-best choice is in italics.

Another way to compare different mortality models is to analyze the standardized residuals, which can be calculated as follows (see Wen et al. 2021):

$$Z_{xtj} = \frac{D_{xtj} - E_{xtj}\hat{m}_{xtj}}{\sqrt{E_{xtj}\hat{m}_{xtj}}} \tag{17}$$

where $D_{xtj}$, $E_{xtj}$, and $\hat{m}_{xtj}$ express the same as earlier. The scatter plot and the heatmap can be used to illustrate the values of the residuals.[10] Based on the heatmap, we can decide whether the cohort effect or more factors should be considered. In Figure A2 in the Appendix, the x-axis refers to calendar years, and the y-axis shows the age structure. The more random the residuals are, the more accurate the mortality model will be. The characteristic patterns in the heatmap may justify the consideration of other factors. For each region and both sexes, the ACF and the ACF–CAE models seem the best. However, there are systematic effects in the heatmaps of the standardized residuals along the cohort years regarding all models. This feature shows that we need to consider the cohort effect, but we do not address this extension in this study. If any systematic parts remain in the standardized residuals after quantifying the cohort effect, we can add a further mortality index to the model to improve the goodness of fit. The ACF and the ACF–CAE models are nested, and we can conclude that these augmented mortality models achieve a better fit. We can present the residuals against age, calendar years, and years of birth using scatter plots, which justify that the model extension improves the estimation (see in the Appendix Figure A3 in the case of the ACF model in the Southern Great Plain).

## Forecasting the mortality indices

To forecast the mortality rates, we need to predict the mortality indices of the models (extrapolate the past trends). The projection was prepared for all models and both sexes. We forecast the mortality indices up to 2050 using autoregressive moving average models. We tested the presence of the unit root (with augmented Dickey–Fuller and Kwiatkowski–Phillips–Schmidt–Shin tests), the normality (with Shapiro–Wilk test) and the autocorrelation of the residuals (with Ljung–Box Q-statistic and Breusch–Godfrey test) related to the time series and their models.[11] The choice between the ARIMA models is based on the results of these tests and the values of the information criteria (AIC and BIC). The results of the different test statistics for the original time series and the accepted ARIMA models, see here.

Assuming the independence of the mortality indices, we can examine the best ARIMA models separately. Figure A4 in the Appendix shows the predicted values of the mortality indices and their 80 and 95% confidence intervals based on the most appropriate time series models related to the ACF model. In most cases, it was

[10] We applied the *StMoMo* package of Villegas et al. (2018) to draw these figures.

[11] We used the *stats* (R Core Team 2021), the *tseries*, and the *lmtest* packages for the testing (see https://CRAN.R-project.org/package=tseries and https://CRAN.R-project.org/package=lmtest).

possible to select the same parsimonious ARIMA process for all regions along the different mortality models. However, the individual forecasts are not based on the same ARIMA model for the Lee–Carter model in the case of males (see Table 4). For example, we fit the ARIMA(0,1,0) model without drift in Northern Hungary because we would have otherwise predicted the decline in mortality. Despite the presence of the COVID-19 epidemic, it is reasonable to assume that life expectancy at birth will increase in the coming decades because of possible medical developments.

Table 4

**ARIMA models of the mortality indices**

| The mortality models | $\kappa_{t(j)}^{(1)}$ | $\kappa_{tj}^{(2)}$ | Regions |
|---|---|---|---|
| | Males | | |
| Lee–Carter | ARIMA(0,1,0) with drift, ARIMA(0,1,0), ARIMA(0,1,1) with drift | – | Northern Great Plain, Northern Hungary, other regions |
| Carter–Lee | ARIMA(0,1,1) with drift | – | all regions |
| CF | ARIMA(0,1,1) with drift | – | all regions |
| CAE | ARIMA(0,1,1) with drift | – | all regions |
| ACF | ARIMA(0,1,0) | ARIMA(0,1,1) with drift | all regions |
| ACF–CAE | ARIMA(0,1,0) | ARIMA(0,1,1) with drift | all regions |
| | Females | | |
| Lee–Carter | ARIMA(0,1,1) with drift | – | all regions |
| Carter–Lee | ARIMA(0,1,1) with drift | – | all regions |
| CF | ARIMA(0,1,1) with drift | – | all regions |
| CAE | ARIMA(0,1,1) with drift | – | all regions |
| ACF | ARIMA(0,1,0) with drift | ARIMA(0,1,1) with drift | all regions |
| ACF–CAE | ARIMA(0,1,0) with drift | ARIMA(0,1,1) with drift | all regions |

During the projection, the values of the mortality indices for 2020 and 2021 were treated as outliers. The number of deaths increased from 2019 to 2020 and from 2020 to 2021 due to the impact of the COVID-19 epidemic. We prepared forecasts following two scenarios expecting higher mortality in 2022 and without future excess deaths. We applied year dummies to predict excess mortality in 2022 and filter out the effect of COVID-19 in 2020–2021. Therefore, we fit and forecast the ARIMA models with an extra variable. The figures of the other models' projected mortality indices and the results of the scenario with excess mortality in 2022, see here. During the calculation of the future mortality rates, we used the predicted values of the mortality indices and the fitted age-varying parameters $\alpha$ and $\beta$. We set the error terms to zero.

## The future life expectancy at birth by region

Future mortality rates can be used to estimate life expectancy at birth. Figure A5 in the Appendix shows the difference in life expectancy at birth[12] over the next 30 years by fitting different mortality models. In the long run, we can assume that future life expectancy will return to the level before the COVID-19 pandemic and that it will continue (assuming that other similar pandemic periods will not occur). Therefore, based on the results of the projection, we can evaluate whether life expectancy at birth reaches or exceeds the level that was observed before the pandemic. From this point of view, we select 2019 as a reference year for the comparison.

The Carter–Lee and CF models do not take into account the regional characteristics of the long-term trend of mortality. Thus, the values for future life expectancy at birth among regions are closer to each other than those estimated by the other models. These two mortality models estimate the lowest life expectancy for men in 2050. The Lee–Carter and the CAE models result in higher values. Overall, the most optimistic forecast is provided by the ACF and ACF–CAE models for males. According to these augmented mortality models, longevity will increase measurably in Budapest, Central Transdanubia, the Northern Great Plain, Pest, the Southern Great Plain, and Southern Transdanubia until 2050. Based on the models with two factors, we predict the following life expectancy at birth for males in 2050 in the abovementioned order of the regions: 78.3–78.9, 74.6–74.6, 74.7–74.7, 75.4–75.9, 74.7–74.8, and 74.8–74.8 years. With the greatest optimism, the ACF–CAE and the CF models forecast 72.1–72.3 years for men at the end of the time horizon in Northern Hungary. Life expectancy at birth will be 76.4–76.5 years for men in 2050 in Western Transdanubia according to the Lee–Carter and the CAE models – the rest of the models estimate a shorter life expectancy.

Analyzing women, we reached a different conclusion – the most favorable forecast is not provided by the ACF and the ACF–CAE models. Generally, the CAE model is the most optimistic, and the augmented models are the most pessimistic. The CAE and the Lee–Carter mortality models predict the highest life expectancy at birth for females in 2050 in Budapest, Central Transdanubia, Pest, Southern, and Western Transdanubia – the lifespan will be 84.0–84.2, 82.8–82.8, 82.5–82.7, 82.7–82.9, and 83.8–83.9 years, respectively. According to the Carter–Lee and CF models, life expectancy will rise to 82.0–82.4 and 81.0–82.4 years by 2050 in the Northern Great Plain and Northern Hungary, respectively. The Carter–Lee and CAE models result in 82.6–82.7 years as the maximum future life expectancy for females in the Southern Great Plain.

[12] During the calculation of the life tables, we follow the methodology of Chiang (1968) and Eurostat (see https://ec.europa.eu/eurostat/cache/metadata/Annexes/demo_mor_esms_an1.pdf), except that the group of people aged 90 and over was considered as the oldest cohort instead of people aged 85 and over.

Three of the six models indicate that male longevity reaches at least the level of 2019 in all regions. In the case of the ACF and the ACF–CAE models, this is likely to occur in 2035–2036, while according to the CAE model, it is likely to occur in 2048. For the other three models, we find at least two regions where the level of life expectancy does not reach the value of 2019 even by 2050. The results are different for women – future life expectancy will exceed the value observed in 2019 after a few years in all regions according to each model. The earliest that this will happen is in 2025 or 2027, which is predicted by the Lee–Carter and the CAE models. Based on the estimation of the ACF–CAE model, women's life expectancy will exceed the value before the COVID-19 pandemic in 2033 in all regions.

Another interesting aspect is how the rank of the regions will change in terms of estimated life expectancy at birth. Previously, we expected that significant rearrangements would not occur in the future. According to the average ranking in 2010–2021, life expectancy was the highest for both men and women in Budapest, while it was the lowest in Northern Hungary. For males, the Northern Great Plain, Pest, and Western Transdanubia followed Budapest in the same rank, while Central Transdanubia, the Southern Great Plain, and Southern Transdanubia had a higher life expectancy than Northern Hungary. The second highest life expectancy for females was observed in Western Transdanubia between 2010 and 2021. This region was followed by Central Transdanubia and Pest, and the Southern Great Plain came in fifth. The Northern Great Plain and Southern Transdanubia overtook Northern Hungary for women.

Regarding males, the Northern Great Plain scored low (7$^{th}$ place) in the ranking of regions based on the Lee–Carter model's forecast. According to the Carter–Lee model's result of placing, the 6$^{th}$ rank of Pest seems less realistic in the next 10 years. The CF model has a similar conclusion for Pest, placing it in 7$^{th}$ place until 2049. In terms of the CAE model, we can argue with the position loss of the Northern Great Plain in the future. In the final years of the forecast horizon, the decline of the Northern Great Plain and the improvement of Southern Transdanubia contradict the trend observed in recent years based on the ACF and the ACF–CAE models.

Analyzing the results of females, we have different conclusions than for men. The ranking based on the Lee–Carter model's prediction seems to be the most reasonable – there will be no major rearrangement in the future. The Carter–Lee model predicts a noticeable loss of position for Pest in the long term, while it results in an improvement for Southern Transdanubia. The Southern Great Plain will also be in a more favorable position than we would expect in advance. The replacement of Budapest and Western Transdanubia is also less realistic. Southern Transdanubia, not Northern Hungary, has the worst position as estimated by the CF model. Based on our prior knowledge, we would not expect the reduction of the rank in Pest, in contrast with the CAE model's projection. In Southern Transdanubia, life expectancy at birth will be higher in the long term than in the period between 2010 and 2021.

This region will be ranked 3rd based on the CAE model. We have similar conclusions regarding the ACF and the ACF–CAE models as the Carter–Lee and the CAE models. Pest will fall to 6th place, and South Transdanubia will be in 4th place, though we find this projection to be less realistic.

We can analyze the differences in future life expectancy at birth between males and females projected by the six mortality models. If we highlight 2019, we can conclude that Budapest and the Northern Great Plain had the slightest difference between the life expectancy of men and women (5.2 and 5.5 years). Furthermore, the most considerable difference was observed in Central Transdanubia and Northern Hungary (7.1 and 7.6 years). The ACF and the ACF–CAE models produce less or nearly the same difference in the long term than was typical in 2019. The predicted gap between males and females narrows further in Budapest regarding the ACF and the ACF–CAE models, in Pest based on the ACF model and in Western Transdanubia according to the ACF–CAE model. All other models are more likely to estimate greater differences in each region, meaning that women's life expectancy will increase compared to that of men.

Checking the models' goodness of fit earlier, we concluded that the ACF and the ACF–CAE models are the most accurate. These findings are also supported by the out-of-sample analysis. Using the data of the period from 1970 to 2015, we prepared forecasts for the years 2016–2019. Based on the comparison of estimated and current mortality rates, the ACF and the ACF–CAE models performed the best for both men and women in almost every region. Regarding the MAPE values, the Lee–Carter model (instead of the ACF–CAE) was the most accurate in addition to the ACF in two cases (for men in Pest and women in Budapest). See more results here.

When comparing multi-population mortality models, an important aspect is to examine whether the forecast produces a coherent result. If the age-specific mortality rates are compared to each other by region and sex, then we can decide whether a given model provides a coherent forecast. If the region-to-region ratios do not diverge in the long run, it means that the differences between the subpopulations are unchanged. Li (2013) and Li et al. (2016) also applied such ratios to identify coherent models. Our analysis confirms that the CF model is clearly coherent. See more results here. According to the region-to-region ratios, the ACF-CAE and the CAE models meet this criterion in the case of males aged above 45 and females aged above 40. Regarding the models with common age effect(s), Wen et al. (2021) found coherent results. The ACF model provided a coherent forecast in the study of Scognamiglio (2022). This is not confirmed by our forecast at all ages. The Lee–Carter and the Carter–Lee models cause divergence problems between the mortality rates of the regions – these findings correspond to our preliminary expectations. The aim of Enchev et al. (2017) was to provide coherent predictions regarding the multi-population models analyzed in their study. They achieved nondivergent forecasts by applying multivariate time series processes.

Analyzing the forecast results, we find that the Lee–Carter model can be appropriate for women if we look at the level of future life expectancy and the long-term evolution of the ranking by region. From the same viewpoints, the mortality models with two factors (the ACF and the ACF–CAE) seem to produce the best results for men. If we assume that the difference in life expectancy between men and women will not increase in the future, then the two-factor models can also be good predictors. This is mainly because these two models are the most pessimistic regarding the life expectancy of women, and we find them among the most optimistic models regarding men. Overall, the augmented models perform well in all aspects for males. For women, the Lee–Carter model seems more realistic when analyzing by region. If we would like to increase the level of the projected life expectancy, the values of the parameters $\alpha_{xj}$ – which are determined based on the data of the entire base period – can be replaced with one of the recent years' (e.g., before the appearance of the COVID-19 pandemic) age-specific log mortality rates. In this way, we can correct the jump-off level, according to Lee–Miller (2001).

Coelho–Nunes (2011) introduce structural break tests following Harvey et al. (2009) and Harris et al. (2009) and forecast the mortality index of the Lee–Carter model with and without an allowance for a structural break. They analyzed different countries and found that structural breaks have been identified more often in data of men than in the case of women. If mortality improvement accelerates after the date of the structural break, then higher life expectancy can be predicted considering the change in the trend. Coelho–Nunes (2011) focus on identifying one structural break, but Sobreira–Nunes (2016) provide tests for multiple breaks in the time series. Presumably, we would also get different projections with higher life expectancy, especially for the Hungarian males, if we followed the methodology of Coelho–Nunes (2011) and Sobreira–Nunes (2016).

## Conclusions

In this study, we focused on some multi-population mortality models that can be considered variants of the original Lee–Carter model. In another study written in 1992, Carter and Lee propose to analyze and predict the common mortality trends of men and women with age-specific coefficients. However, this modification does not necessarily guarantee a coherent forecast. Li–Lee (2005) already mention mortality models for which divergence does not arise as a problem in the long run. If the model is calculated with the same mortality index values and coefficient for all subpopulations, it will eliminate the divergence. With an additional time-varying factor, the periodic difference from the long-term trend can be incorporated into the model, which can be different for the groups. The age-dependent coefficient of the second mortality index may even be the same regarding the subpopulations. The latter variant of the multi-population models is mentioned by Li et al. (2016) following the idea of Kleinow (2015). We can modify the Lee–Carter model with a common age

effect that allows the mortality index to vary between the groups, as suggested by Kleinow (2015).

We implemented the Poisson Lee–Carter and five multi-population mortality models to predict life expectancy in Hungary by region. Comparing the goodness of fit between the different models, the ACF and the ACF–CAE models proved to be the best choice for both men and women. These models also performed well for men regarding the remaining long-term mortality differences between the regions. Furthermore, the augmented models reduced or kept the life expectancy gap between men and women at the same level by region. However, the values of future life expectancy at birth appear low for females based on the ACF and ACF–CAE models. For women, the Lee–Carter model would be preferred from this point of view. It can be the case that a subgroup does not follow the patterns of the whole population, but we expect that the subgroup will follow the whole population's mortality level in the future. Therefore, using the common mortality index does not necessarily have to be excluded during the projection even if there are significant differences between the groups based on the current data. These models can also provide a possible scenario for future life expectancy.

To improve the models' goodness of fit, a possible development is to estimate the models with cohort effect, as Yang et al. (2016) examined. They consider both the group-specific and common cohort effect in multi-population mortality models. In addition to regional modeling, we can develop a forecast for counties using common parameters ($\beta$ and/or $\kappa$) by region. According to the explanation ratio, it is possible to decide what is the most appropriate for the counties, for example, using the national or regional mortality trend. Based on the comparison of this ratio, Li–Lee (2005) suggest analyzing whether it is necessary to exclude a subgroup from a population. If a subpopulation is excluded from a greater group, then the common and population-specific factors must be re-estimated, just as they would when entering a group. It could be the subject of further analysis to cluster the parameter $\beta$ by age groups, as suggested by Kleinow (2015). Regarding these cohorts, we could use the common age effect for the subpopulations. A further extension could be if the subgroups were clustered on the basis of the parameter $\beta$, so these values would be the same within some greater groups, but different between them. It is also possible to create larger groups based on the similarity of the group-specific mortality indices, which would be common for the population members.

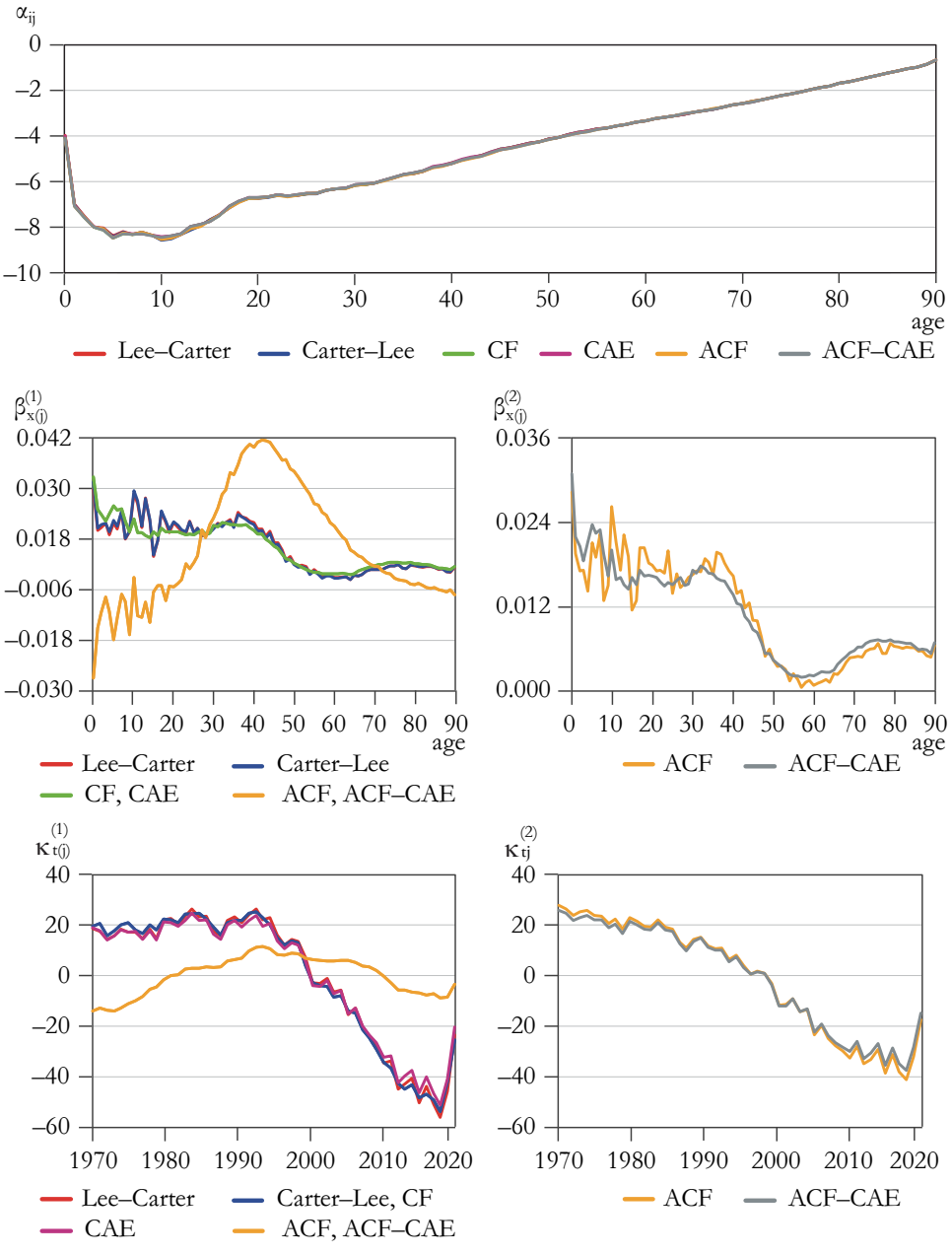
## Acknowledgments

I would like to thank Dávid Burka and László Kovács for their valuable comments. I am also grateful to the Cooperative Doctoral Programme of the Ministry of Culture and Innovation financed by the National Research, Development and Innovation Fund and the Hungarian Central Statistical Office for the opportunity to analyze Hungarian regional data.

## APPENDIX

Figure A1

The fitted parameters of the mortality models in the Southern Great Plain Males




*(Figure continues on the next page.)*

*(Continued.)*

**Females**

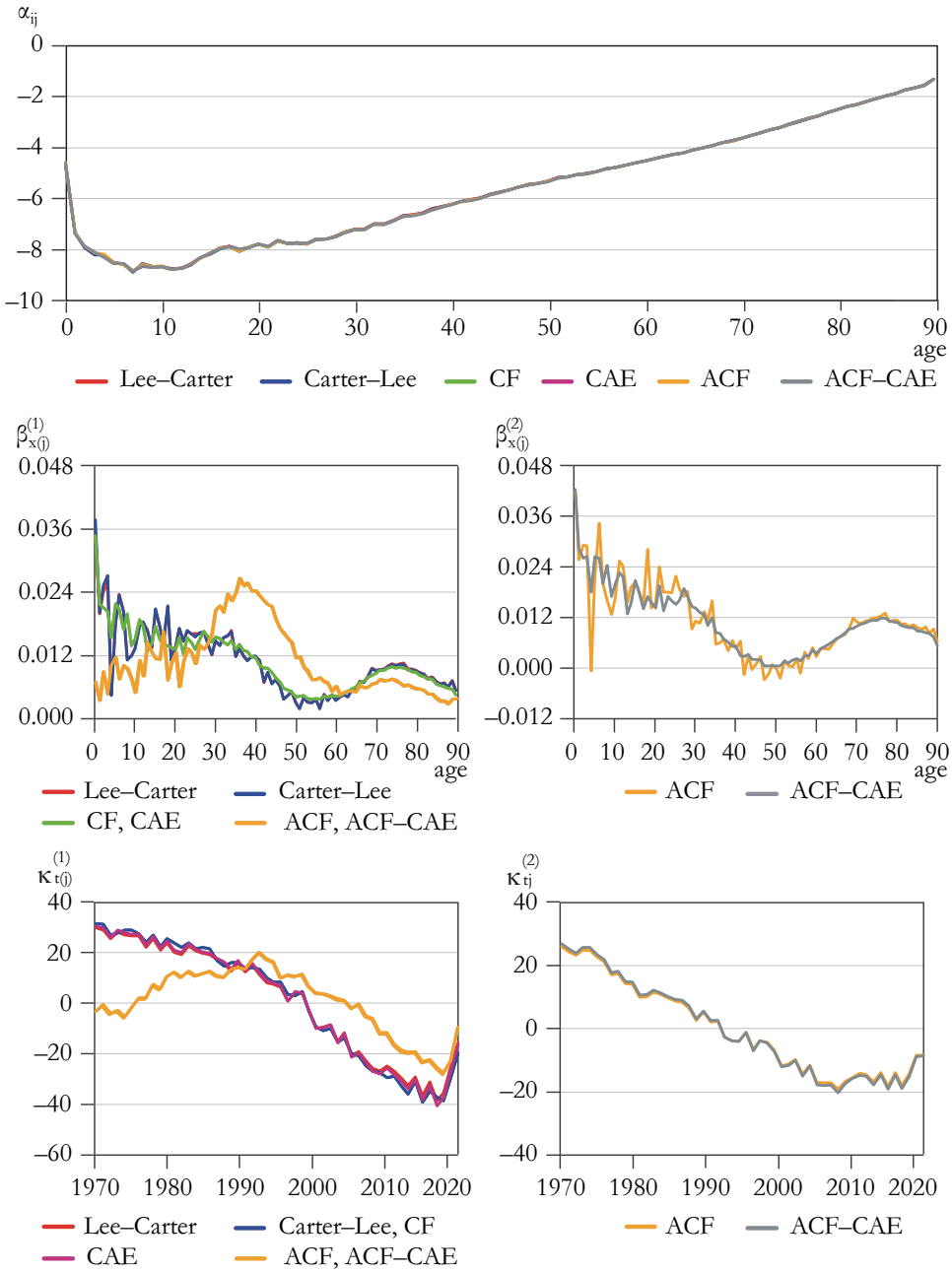
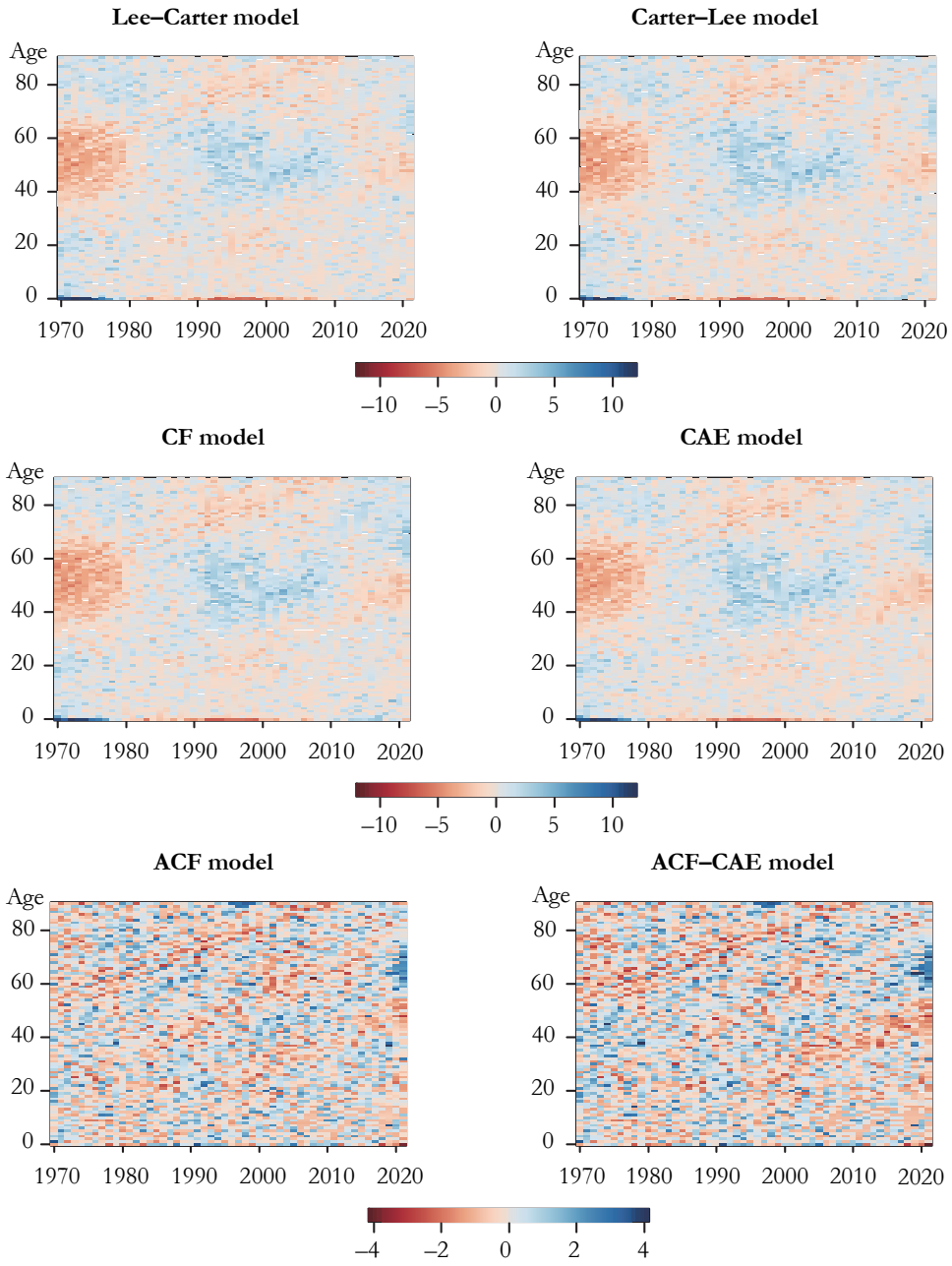


Figure A2

**The standardized residuals of the mortality models in the Southern Great Plain Males**




*(Figure continues on the next page.)*

*(Continued.)*

**Females**

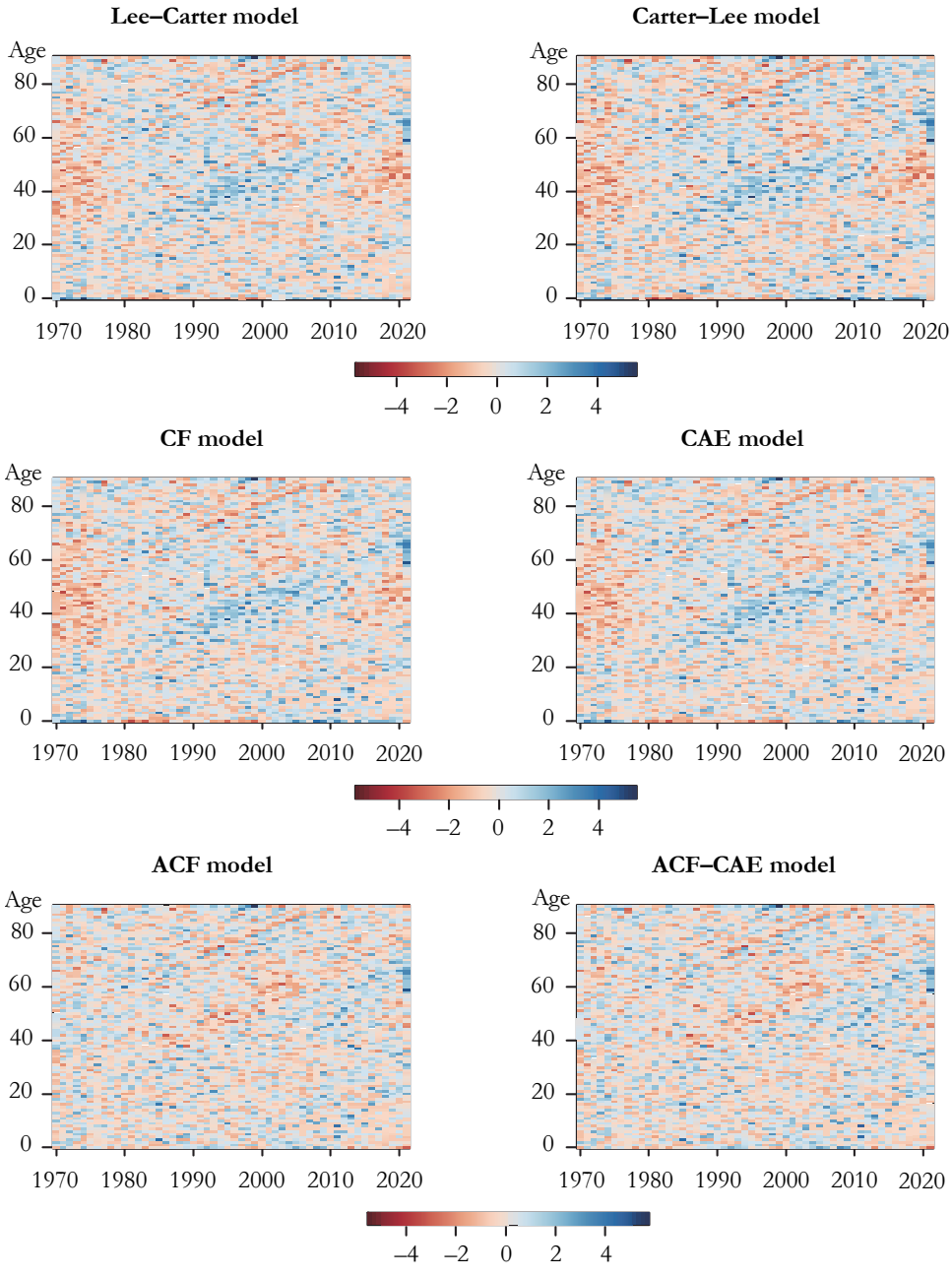


Figure A3

The standardized residuals of the ACF model in the Southern Great Plain

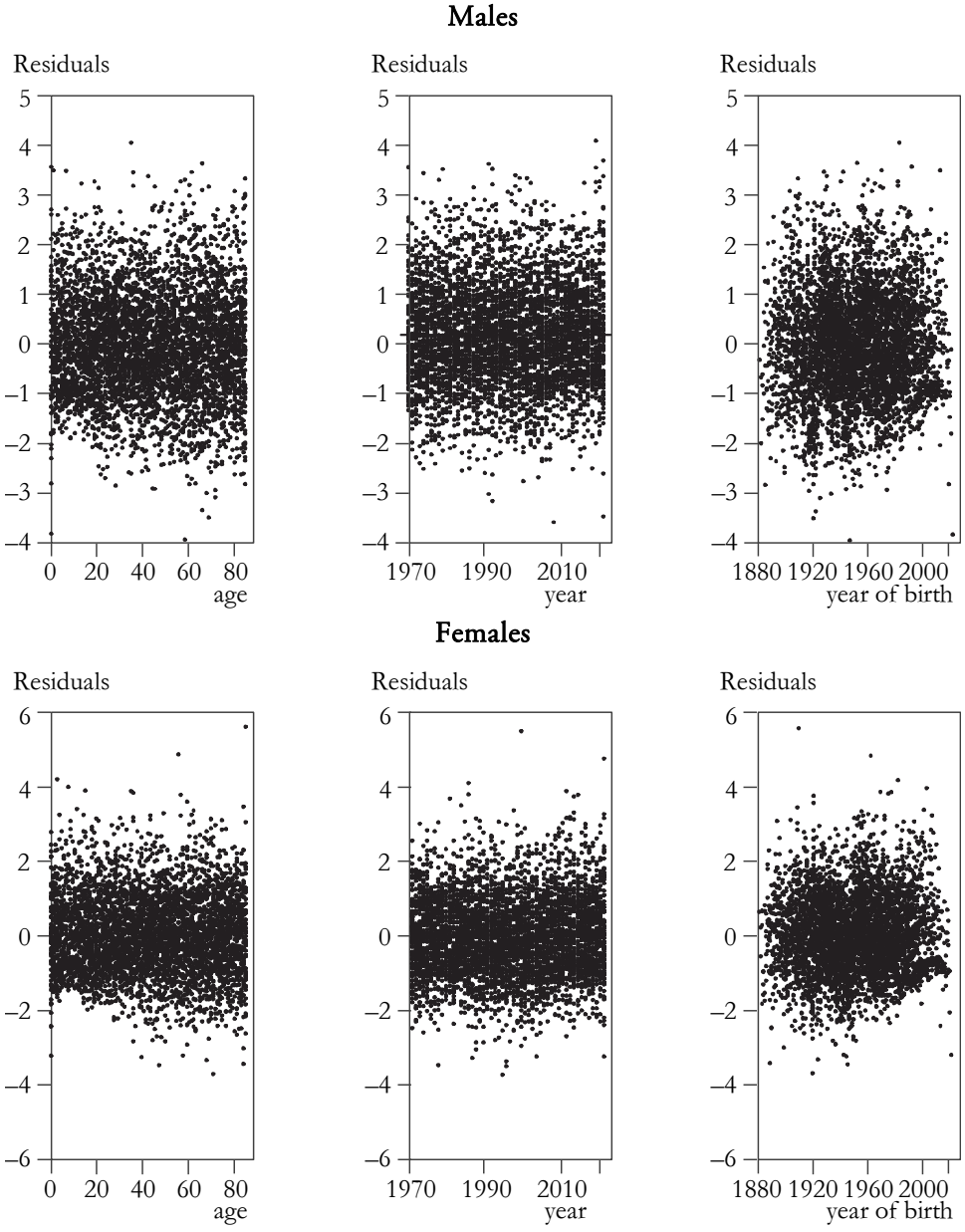
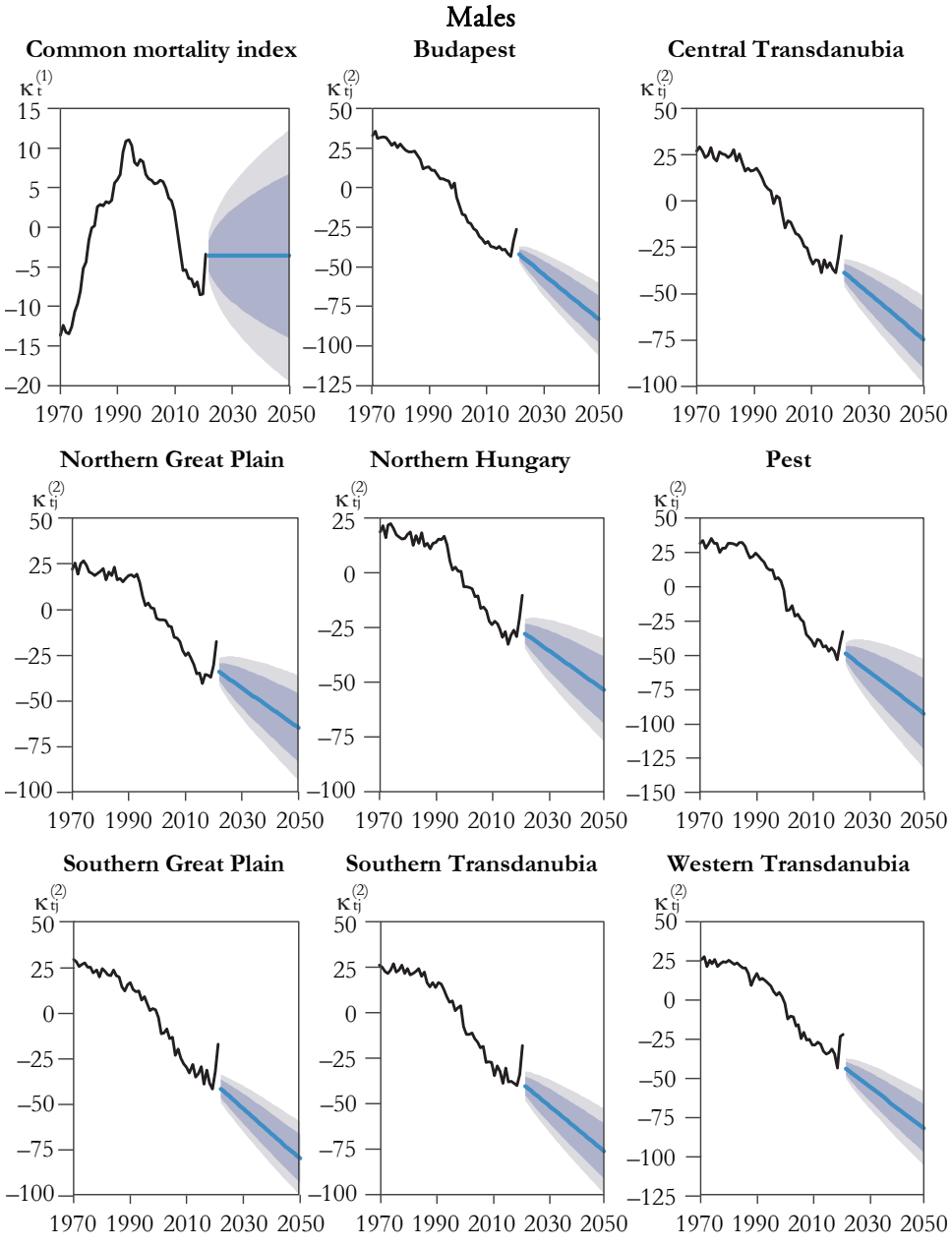


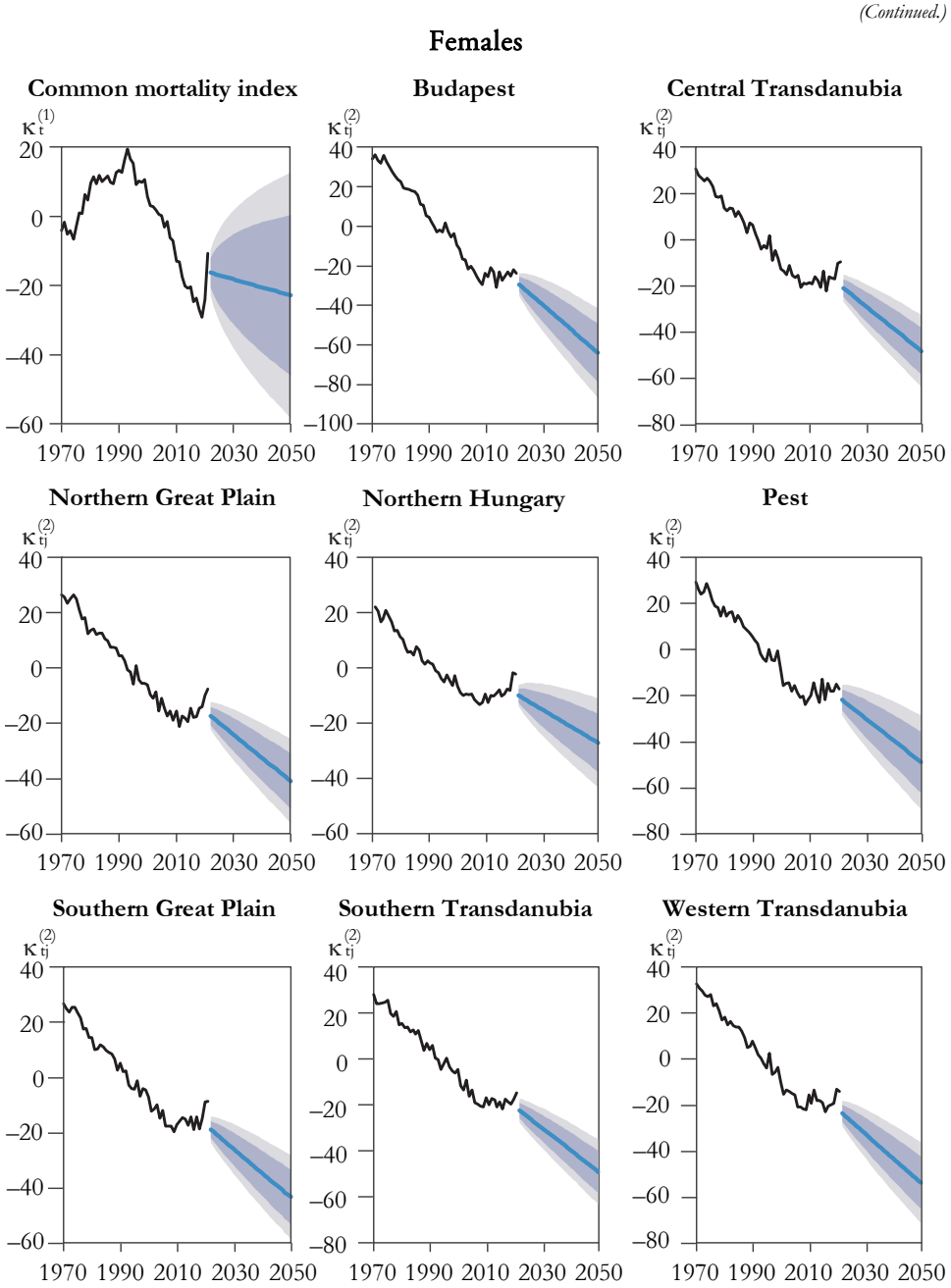
Figure A4

**The common and group-specific mortality indices of the ACF model**




*(Figure continues on the next page.)*

*(Continued.)*

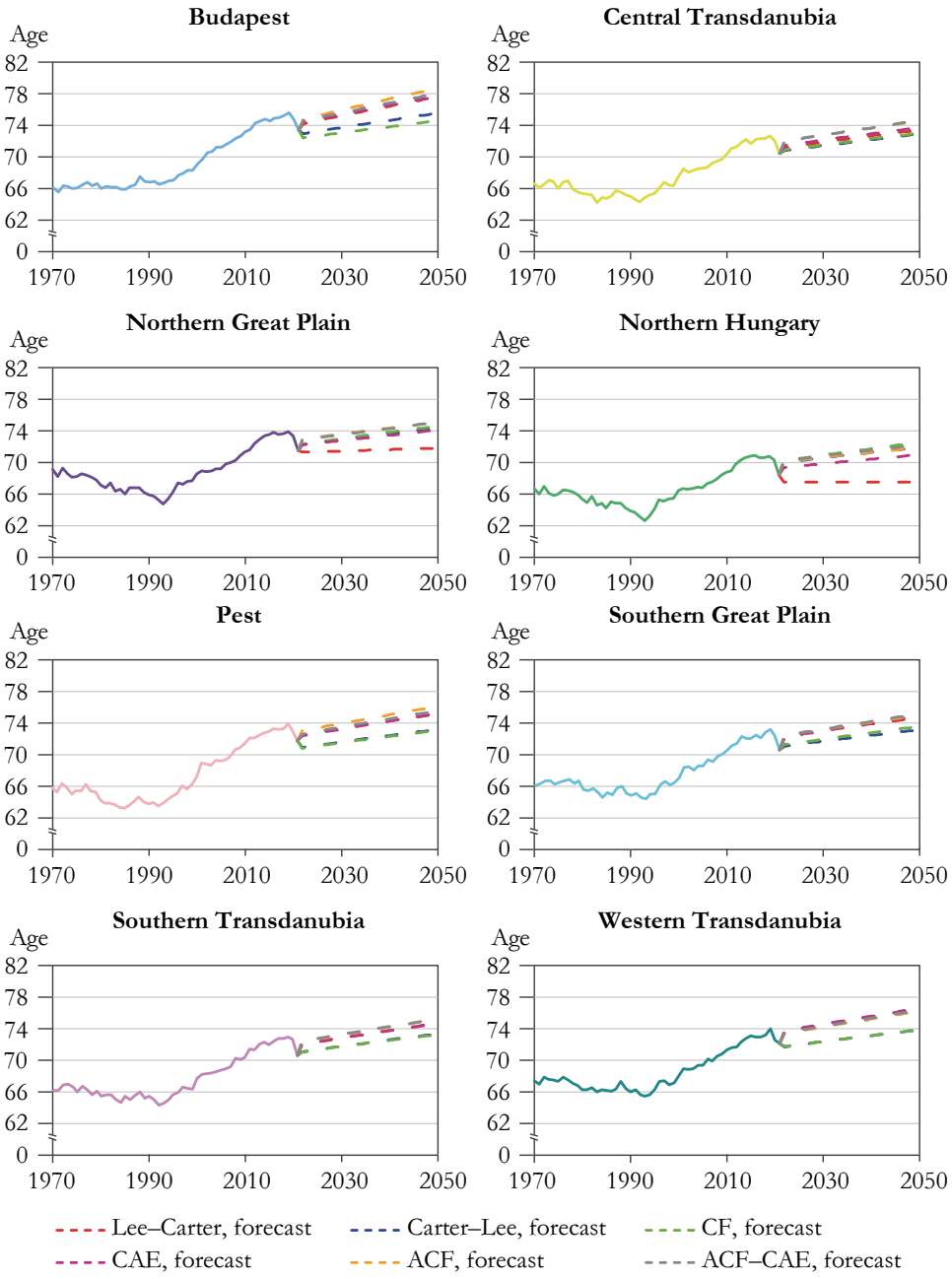



*Note:* The gray backgrounds represent the 80 and 95% confidence intervals.

Figure A5

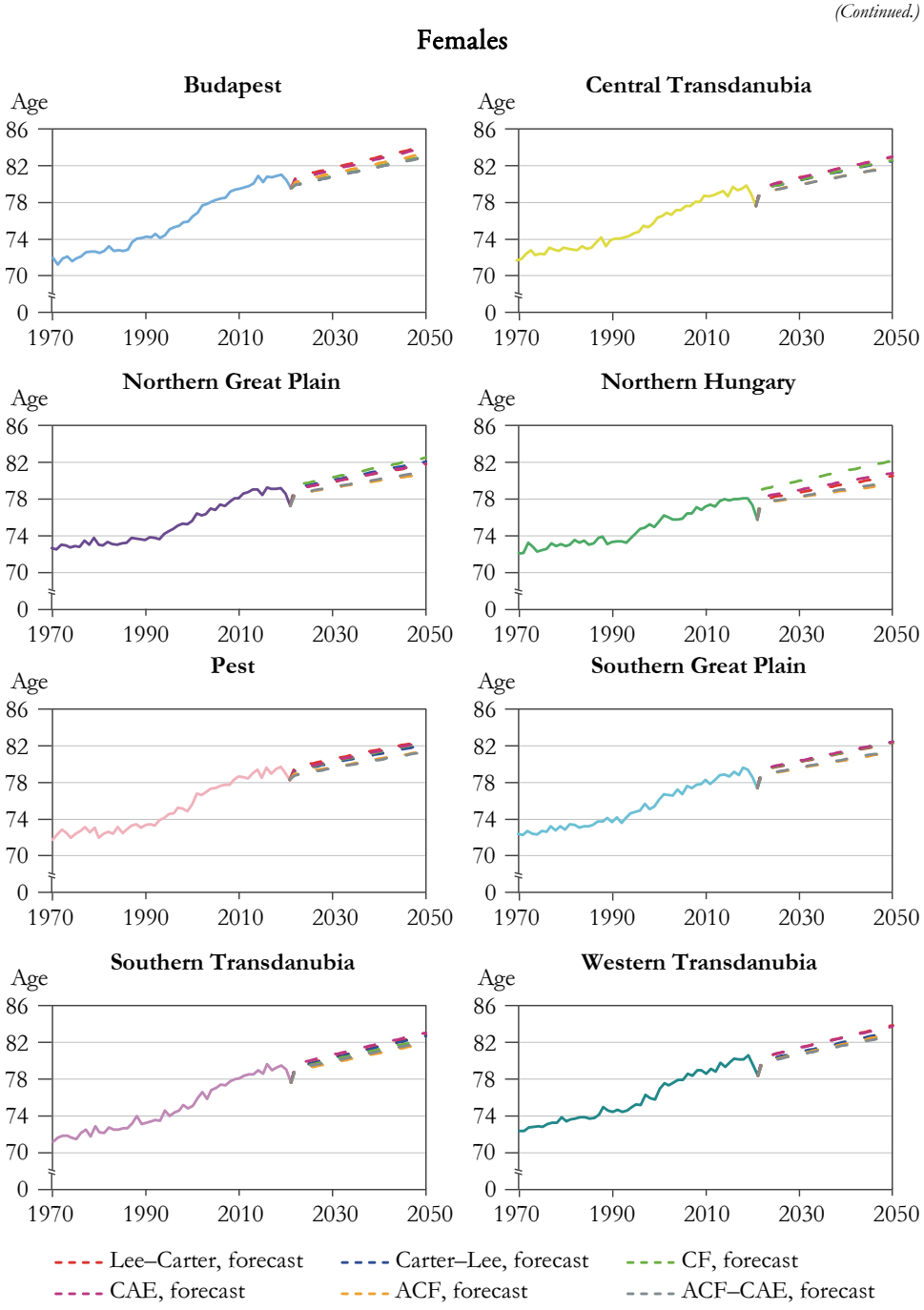
**The current and future life expectancy at birth by region**

**Males**




*(Figure continues on the next page.)*

*(Continued.)*


Females
Budapest
Central Transdanubia
Northern Great Plain
Northern Hungary
Pest
Southern Great Plain
Southern Transdanubia
Western Transdanubia
Age
86
82
78
74
70
0
1970
1990
2010
2030
2050
Lee–Carter, forecast
Carter–Lee, forecast
CF, forecast
CAE, forecast
ACF, forecast
ACF–CAE, forecast

## REFERENCES


BÁLINT, L. (2016): Mennyire illeszkedik a magyar halandóság alakulása az epidemiológiai átmenet elméletéhez? *Demográfia* 59 (1): 5–57. https://doi.org/10.21543/Dem.59.1.1

BRILLINGER, D. R. (1986): The natural variability of vital rates and associated statistics *Biometrics* 42 (4): 693–734. https://doi.org/10.2307/2530689

BROUHNS, N.–DENUIT, M.–VERMUNT, J. K. (2002): A Poisson log-bilinear regression approach to the construction of projected lifetables *Insurance: Mathematics and Economics* 31 (3): 373–393. https://doi.org/10.1016/S0167-6687(02)00185-3

CARTER, L. R.–LEE, R. D. (1992): Modeling and forecasting US sex differentials in mortality *International Journal of Forecasting* 8 (3): 393–411. https://doi.org/10.1016/0169-2070(92)90055-E

CHIANG, C. L. (1968): *Introduction to stochastic processes in biostatistics* Wiley, New York.

COELHO, E.–NUNES, L. C. (2011): Forecasting mortality in the event of a structural change *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 174 (3): 713–736. https://doi.org/10.1111/j.1467-985X.2010.00687.x

DANESI, I. L.–HABERMAN, S.–MILLOSSOVICH, P. (2015): Forecasting mortality in subpopulations using Lee–Carter type models: A comparison *Insurance: Mathematics and Economics* 62: 151–161. http://dx.doi.org/10.1016/j.insmatheco.2015.03.010

ENCHEV, V.–KLEINOW, T.–CAIRNS, A. J. G. (2017): Multi-population mortality models: Fitting, forecasting and comparisons *Scandinavian Actuarial Journal* 2017 (4): 319–342. https://doi.org/10.1080/03461238.2015.1133450

HARRIS, D.–HARVEY, D. I.–LEYBOURNE, S. J.–TAYLOR, A. M. R. (2009): Testing for a unit root in the presence of a possible break in trend *Econometric Theory* 25 (6): 1545–1588. https://doi.org/10.1017/S0266466609990259

HARVEY, D. I.–LEYBOURNE, S. J.–TAYLOR, A. M. R. (2009): Simple, robust, and powerful tests of the breaking trend hypothesis *Econometric Theory* 25 (4): 995–1029. https://doi.org/10.1017/S0266466608090385

HUNGARIAN CENTRAL STATISTICAL OFFICE (HCSO) (1990): *Korrigált népességszámok megyénként, 1980–1990* Központi Statisztikai Hivatal, Budapest.

HUNGARIAN CENTRAL STATISTICAL OFFICE (HCSO) (2001): *A 2001. február 1-jei népszámlálás végleges adatai alapján korrigált 1990–2001. évi továbbszámított népességszámok. Munkaanyag* Központi Statisztikai Hivatal, Budapest.

HUNGARIAN CENTRAL STATISTICAL OFFICE (HCSO) (2021): *Demographic Yearbook, 2020* Hungarian Central Statistical Office, Budapest.

HUNT, A.–BLAKE, D. (2014): A general procedure for constructing mortality models *North American Actuarial Journal* 18 (1): 116–138. https://doi.org/10.1080/10920277.2013.852963

JÓZAN, P. (2008): *Válság és megújulás a második világháború utáni epidemiológiai fejlődésben Magyarországon* MTA Társadalomkutató Központ, Budapest.

JÓZAN, P. (2009): Csökkenő kardiovaszkuláris mortalitás, javuló életkilátások, új epidemiológiai korszak kezdete Magyarországon *IME Interdiszciplináris Magyar Egészségügy* 8 (4): 21–25.

JÓZAN, P. (2012): Rendszerváltozás és epidemiológiai korszakváltás Magyarországon *Orvosi Hetilap* 153 (17): 662–677. https://doi.org/10.1556/oh.2012.29344

KAPÁS, J. (2022): Has COVID-19 caused a change in the dynamics of the unemployment rate? The case of North America and continental Europe *Regional Statistics* 12 (1): 3–26. https://doi.org/10.15196/RS120107

KHEDHIRI, S. (2022): Comparison of SARFIMA and LSTM methods to model and to forecast Canadian temperature *Regional Statistics* 12 (2): 177–194. https://doi.org/10.15196/RS120204

KINCSES, Á.–TÓTH, G. (2020): How coronavirus spread in Europe over time: national probabilities based on migration networks *Regional Statistics* 10 (2): 228–231. https://doi.org/10.15196/RS100210

KLEINOW, T. (2015): A common age effect model for the mortality of multiple populations *Insurance: Mathematics and Economics* 63: 147–152. https://doi.org/10.1016/j.insmatheco.2015.03.023

LEE, R. D. (2000): The Lee-Carter method for forecasting mortality, with various extensions and applications *North American Actuarial Journal* 4 (1): 80–93. https://doi.org/10.1080/10920277.2000.10595882

LEE, R. D.–CARTER, L. R. (1992): Modeling and forecasting U.S. mortality *Journal of the American Statistical Association* 87 (419): 659–671. https://doi.org/10.2307/2290201

LEE, R. D.–MILLER, T. (2001): Evaluating the performance of the Lee-Carter method for forecasting mortality *Demography* 38 (4): 537–549. https://doi.org/10.1353/dem.2001.0036

LEE, R. D.–NAULT, F. (1993): *Modeling and forecasting provincial mortality in Canada* Presented at the World Congress of the International Union for Scientific Study of Population. Montréal, Canada.

LI, J. (2013): A Poisson common factor model for projecting mortality and life expectancy jointly for females and males *Population Studies* 67 (1): 111–126. https://doi.org/10.1080/00324728.2012.689316

LI, J.–TICKLE, L.–PARR, N. (2016): A multi-population evaluation of the Poisson common factor model for projecting mortality jointly for both sexes *Journal of Population Research* 33 (4): 333–360. https://doi.org/10.1007/s12546-016-9173-0

LI, N.–LEE, R. D. (2005): Coherent mortality forecasts for a group of populations: An extension of the Lee–Carter method *Demography* 42 (3): 575–594. https://doi.org/10.1353%2Fdem.2005.0021

RENSHAW, A.–HABERMAN, S. (2006): A cohort-based extension to the Lee–Carter model for mortality reduction factors *Insurance: Mathematics and Economics* 38 (3): 556–570. https://doi.org/10.1016/j.insmatheco.2005.12.001

SCOGNAMIGLIO, S. (2022): Longevity risk analysis: applications to the Italian regional data *Quantitative Finance and Economics* 6 (1): 138–157. https://doi.org/10.3934/QFE.2022006

SOBREIRA, N.–NUNES, L. C. (2016): Tests for multiple breaks in the trend with stationary or integrated shocks *Oxford Bulletin of Economics and Statistics* 78 (3): 394–411. https://doi.org/10.1111/obes.12116

VILLEGAS, A. M.–HABERMAN, S.–KAISHEV, V. K.–MILLOSSOVICH, P. (2017): A comparative study of two-population models for the assessment of basis risk in longevity hedges *ASTIN Bulletin* 47 (3): 631–679. https://doi.org/10.1017/asb.2017.18

VILLEGAS, A. M.–KAISHEV, V. K.–MILLOSSOVICH, P. (2018): StMoMo: An R package for stochastic mortality modeling *Journal of Statistical Software* 84 (3): 1–38. https://doi.org/10.18637/jss.v084.i03

WEN, J.–CAIRNS, A. J. G.–KLEINOW, T. (2021): Fitting multi-population models to socio-economic groups *Annals of Actuarial Science* 15 (1): 144–172. https://doi.org/10.1017/S1748499520000184

YANG, B.–LI, J.–BALASOORIYA, U. (2016): Cohort extension of the Poisson common factor model for modelling both genders jointly *Scandinavian Actuarial Journal* 2016 (2): 93–112. https://doi.org/10.1080/03461238.2014.908411

## INTERNET SOURCES

BÁLINT, L. (2011): A születéskor várható élettartam nemek szerinti térbeli különbségei *Területi Statisztika* 51 (4): 386–404. https://www.ksh.hu/statszemle_archive/terstat/2011/2011_04/ts2011_04_04.pdf (downloaded: October 2023)

EUROSTAT: *Description of the Eurostat method for the calculation of the life expectancies at all ages.* https://ec.europa.eu/eurostat/cache/metadata/Annexes/demo_mor_esms_an1.pdf (downloaded: October 2022)

HUNT, A.–BLAKE, D. (2015): On the structure and classification of mortality models *Pensions Institute Working Paper* PI-1506 http://www.pensions-institute.org/wp-content/uploads/2019/workingpapers/wp1506.pdf (downloaded: October 2023)

R CORE TEAM (2021): *R: A language and environment for statistical computing* R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/ (downloaded: October 2022)

TRAPLETTI, A.–HORNIK, K. (2022): *Time Series Analysis and Computational Finance. R package version 0.10-52.* https://CRAN.R-project.org/package=tseries (downloaded: October 2022)

ZEILEIS, A.–HOTHORN, T. (2002): Diagnostic Checking in Regression Relationships *R News* 2 (3): 7–10. https://CRAN.R-project.org/doc/Rnews/ and https://CRAN.R-project.org/package=lmtest (downloaded: October 2022)