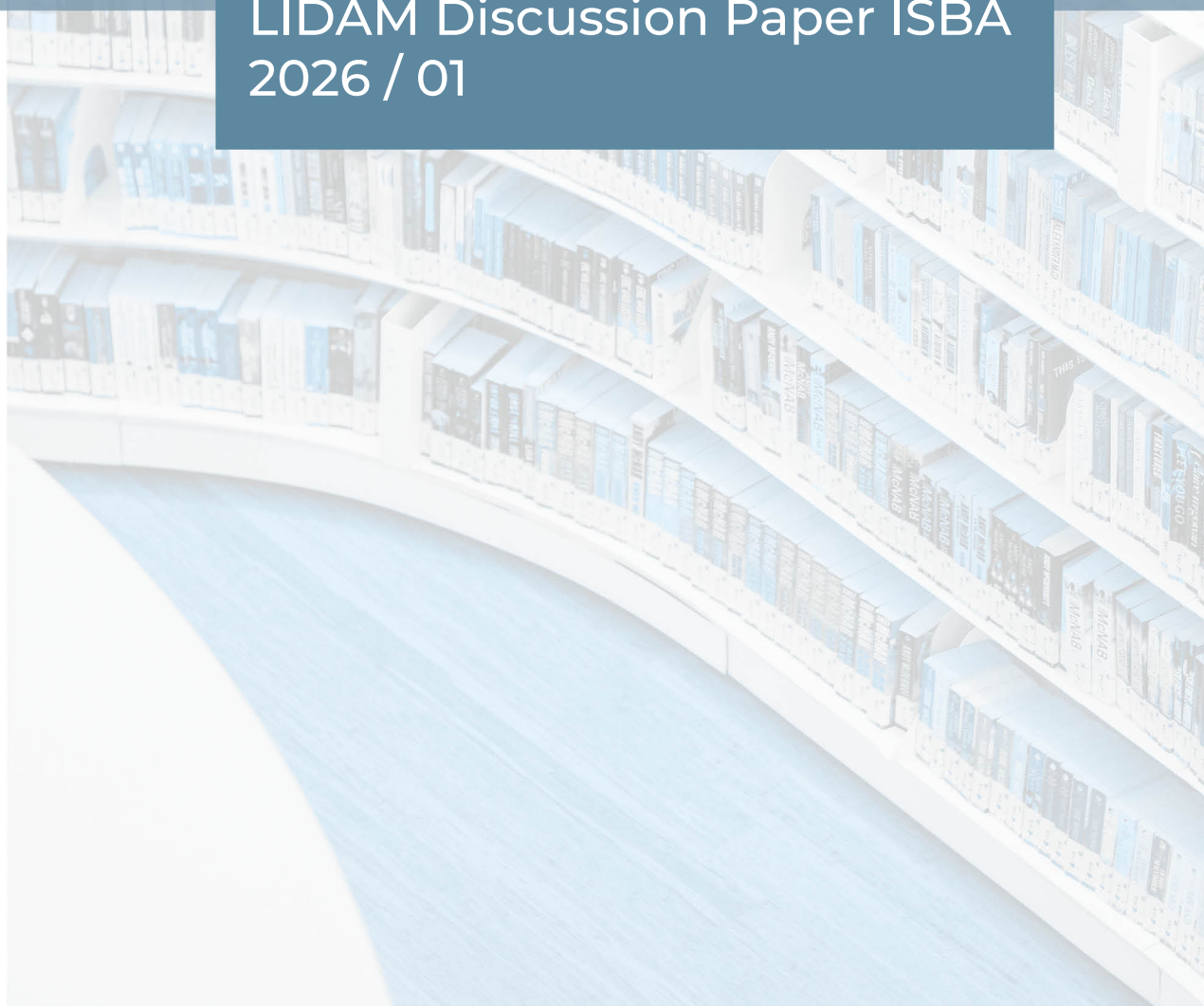
AN ECONOMIC-ENVIRONMENTAL APPROACH FOR REGIONAL MORTALITY

Donatien Hainaut

LIDAM Discussion Paper ISBA
2026 / 01



UCLouvain

ISBA | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/isba/publication

An economic-environmental approach for regional mortality

Donatien Hainaut*
UCLouvain - LIDAM

This article proposes a parametric model explaining differences in mortality across European NUTS2 regions using economic and environmental variables. We extend a multi-group version of the Lee–Carter framework by incorporating economic and environmental time series as driving factors. The marginal effects of these factors are modeled with B-splines. Compared to the Li and Lee framework, our model offers several advantages. First, it is interpretable and allows assessment of the impact on mortality resulting from changes in economic or environmental policies. Second, the parameterization limits the model's degrees of freedom, enabling reliable estimation over shorter time windows. We illustrate the efficiency of our approach by explaining regional mortality patterns in France, Italy, and Belgium–Netherlands.

KEYWORDS : mortality forecasting, Lee-Carter model, multi-group mortality, life insurance

1 Introduction

The Lee–Carter (LC) model (1992) has become a standard method for forecasting death rates. Since its publication, multiple extensions have been proposed. Without being exhaustive, we cite the age–period–cohort version of Renshaw and Haberman (2006), the Cairns, Blake, and Dowd models (2008), and the extension of Plat (2009). Hyndman et al. (2013) proposed an extension based on functional principal component models of simple functions of death rates. Similarly, Hainaut (2018) introduced an LC variation based on a neural encoder–decoder.

On the other hand, the impact of economic development on mortality has been clearly established in the literature. For instance, Brenner (2005) showed that GDP per capita, over the long term, is inversely related to mortality rates during 1901–2000 in the USA. However, in the very short term, rapid economic growth is occasionally associated with increased mortality rates. Birchenall (2007) finds evidence that improvements in economic conditions since the 18th century are an important factor behind the decline in death rates in Northwestern Europe. Conversely, Tapia Granados and Roux (2009) argue that US mortality declines accelerate during economic recessions and slow down during economic expansions. These ambiguous results are discussed in Ruhm (2005), who highlights the difficulties in adequately controlling for confounding factors that are spuriously correlated with macroeconomic conditions.

*Corresponding author. Postal address: Voie du Roman Pays 20, 1348 Louvain-la-Neuve Belgium. E-mail to: donatien.hainaut(at)uclouvain.be

Motivated by the visible link between mortality and economic conditions, French and O'Hare (2014) incorporated macroeconomic variables into a mortality model and found that these factors improve the predictive capacity of the Lee–Carter model. Boonen and Li (2017) extended the Li and Lee (2005) model by including real gross domestic product (GDP) per capita, combined with a singular value decomposition applied to population-specific residuals of mortality rates not explained by the common Li and Lee factor. Dimai (2024) incorporated variables related to economic, environmental, and lifestyle factors to predict mortality in 37 countries. The model uses principal components derived from these variables, extending the Niu and Melenberg (2014) model beyond GDP. More recently, Villegas et al. (2025) investigated how mortality trends for leading causes of death differ between socio-economic groups.

This article examines whether environmental and economic factors explain differences in regional NUTS2[1] mortality rates. Modeling regional mortality presents several challenges. First, regional death rates are more volatile than national ones due to smaller population sizes. A non-parametric multi-group model, such as Li and Lee (2005), therefore tends to overfit data and yields non-smooth forecasts. Second, mortality exhibits different patterns by region that are not easily explained by models relying solely on endogenous risk factors. The model proposed in this article addresses these two issues. To avoid overfitting, we consider a multi-group parametric extension of the Lee–Carter model based on B-splines. These provide smooth, flexible curves for modeling complex shapes, offering excellent local control and computational efficiency. B-splines have been successfully applied to mortality modeling in recent publications: Camarda (2019) uses P-splines, which combine B-splines with a smoothness penalty, for forecasting death rates; Oirov et al. (2021) approximate Mongolian death rates with B-splines; and Pavone et al. (2024) propose a B-spline mortality model with locally adaptive dynamic coefficients.

In our framework, we use linear combinations of B-splines with fixed knots to model the marginal effect of risk factors by age. This ensures parsimony and smooth forecast curves of death rates. To improve interpretability, the model considers only a single endogenous risk factor with regional sensitivities, coupled with five economic and three environmental factors. The model, called the Economic–Environmental Lee–Carter (Eco-Env LC), relies on relative GDP per inhabitant, relative purchasing power standard per inhabitant, an index of regional education, employment rate, and the number of physicians per hundred thousand inhabitants. The environmental risk factors are regional emissions of NO_2 in kilotonnes of CO_2 equivalent per square kilometer, relative heating degree days (HDD), and cooling degree days (CDD). Compared to existing literature, such as Dimai (2024), our parametric approach requires fewer parameters and does not use principal components as explanatory variables. Compared to Niu and Melenberg (2014) or Boonen and Li (2017), which use only GDP per capita, we include eight external risk factors.

The paper is organized as follows. The next section introduces the Eco-Env LC model and describes the economic and environmental factors. The third section provides details about the parameterization with B-splines. Estimation is performed using an approximated Newton–Raphson method with adaptive learning. In the fourth section, we fit the Eco-Env LC model to regional mortality rates for France (FR), Italy (IT), and Belgium–Netherlands (BE-NL). The goodness of fit is benchmarked against the Lee–Carter (LC), a parametric LC, and the Li and Lee models. Next, we compare the marginal effects of economic and environmental variables on mortality. Finally, we forecast period life expectancies and estimate the impact of changes in economic and environmental factors on these predictions.

[1] NUTS 2 regions are the basic statistical regions within the European Union, typically with a population of 800,000 to 3 million, used for the application of regional policies.

2 Model

We use economic, climatic and mortality data for European NUTS 2 regions from the Eurostat portal[2]. This database is directly accessible from Python via an API, which provides great flexibility of use. For a given country, let $D_{x,t,g}$ denote the number of deaths at age x, during calendar year t, in the g^{th} NUTS 2 region. Depending on the country, the age range extends from 0 to a maximum between 80 and 85 years old. We denote the minimum and maximum ages by x_m and x_M, respectively. The number of calendar years also varies among countries. The time window considered for modeling mortality is constrained by the availability of economic and environmental factors. For illustration, we consider periods from 2001 up to 2024 or 2023, depending on the country.

Because data are available over a restricted time range, we develop a parametric model that is less sensitive to measurement errors in death rates than a non-parametric model. As shown in the numerical illustrations, this approach allows us to explain differences between regional mortality trends. The minimum and maximum calendar years are denoted by t_m and t_M. The number of NUTS 2 regions, denoted g_M, depends on the size of the country. The Eurostat portal also provides exposures $E_{x,t,g}$ and death rates $\mu_{x,t,g}$ by age, calendar year and region. Following Brouhns et al. (2002), we assume that the number of deaths, $D_{x,t,g}$ follows a Poisson distribution with intensity $\mu_{x,t,g}$:

$$D_{x,t,g} \quad \sim \quad Po\left(E_{x,t,g}\mu_{x,t,g}\right) .$$

This assumption makes the model more robust than using a binomial distribution, and the likelihood is proportional to the binomial one. The probability of observing $d_{x,t,g}$ deaths is provided by the standard formula:

$$P\left(D_{x,t,g} = d_{x,t,g}\right) \quad = \quad \frac{\exp\left(-\mu_{x,t,g}\, E_{x,t,g}\right)\left(\mu_{x,t,g}\, E_{x,t,g}\right)^{d_{x,t,g}}}{d_{x,t,g}!} .$$

Factors	Dataset names	Variables
Mortality	demo_r_mlife	DEATHRATE
		PYLIVED
		NUMBERDYING
GDP	nama_10r_2gdp	EUR_HAB_EU27_2020
Households	nama_10r_2hhinc	PPS_EU27_2020_HAB
Education	edat_lfse_04	ED0-2 , T, Y25-64
		ED3_4 , T, Y25-64
		ED5-8 , T, Y25-64
Employment	lfst_r_lfe2emprt	T, Y25-64
Physicians	hlth_rs_physreg	P_HTHAB
Heating degree days	hddidx (NUTS3)	Annual values
Cooling degree days	cddidx (NUTS3)	Annual values
NO2	EDGAR_2025_GHG_NUTS2	N2O

Table 1: List of datasets and variables used by the Eco-Env LC model. Except NO2 data coming from EDGAR, all other statistics come from Eurostat.

We relate the death rates to a set $\mathcal{F}$ of regional economic and environmental factors, denoted by $\kappa_{t,g}^{(f)}$ for $f \in \mathcal{F}$. In our Economic–Environmental Lee–Carter (Eco-Env LC) model, we postulate

[2]https://ec.europa.eu/eurostat/web/main/data/database

the following dynamics for log-death rates:

$$\ln(\mu_{x,t,g}) = \alpha_x + \beta_{x,g}^{(end)}\kappa_t^{(end)} + \sum_{f\in\mathcal{F}} \beta_x^{(f)}\kappa_{t,g}^{(f)}\,, \tag{1}$$

where $\beta_x^{(f)}$ represents the marginal impact of the f^{th} factor on the death rate at age x. Similar to the original Lee–Carter model, α_x is a baseline mortality rate. In addition to the exogenous factors $\mathcal{F}$, we consider a single endogenous factor, $\kappa_t^{(end)}$, common to all NUTS 2 regions. Its marginal impact on mortality, denoted by $\beta_{x,g}^{(end)}$, varies by region to capture residual heterogeneity in death rates.

To limit the number of parameters to estimate, the curves α_x, $\beta_x^{(end)}$ and $\beta_x^{(f)}$ for all $f \in \mathcal{F}$, are expressed as linear combinations of B-splines. The construction of these curves is detailed in the next section. Before that, we introduce the factors in $\mathcal{F}$ as reported in Table 1.

The first economic factor is proportional to the Gross Domestic Product (GDP) per inhabitant, expressed as a percentage of the EU27 (2020) average. We denote it by $GDP_{t,g}$ for $t = t_m$ to t_M and $g = 0$ to g_M. Since our aim is to explain regional differences, we define the economic factor $\kappa_t^{(gdp)}$, as the ratio of $GDP_{t,g}$ to the average $GDP_{t,g}$ across all regions:

$$\kappa_t^{(gdp)} = \frac{GDP_{t,g}}{\frac{1}{g_M}\sum_{g=0}^{g_M} GDP_{t,g}}\,.$$

We denote by $\beta_x^{(gdp)}$ the marginal effect of the GDP factor on mortality at age x.

The second factor relates to the Purchasing Power Standard (PPS, EU27 from 2020), expressed per capita as a percentage of the EU27 (2020) average. This statistic is denoted by $PPS_{t,g}$. As with GDP, we define the corresponding factor, $\kappa_t^{(pps)}$ as the ratio of $PPS_{t,g}$ to the average PPS across all regions:

$$\kappa_t^{(pps)} = \frac{PPS_{t,g}}{\frac{1}{g_M}\sum_{g=0}^{g_M} PPS_{t,g}}\,,$$

The marginal effect of PPS on mortality is denoted by $\beta_x^{(pps)}$.

Next, we define a risk factor related to the regional educational level. For this purpose, we retrieve the percentage of the total population aged 25 to 64 years with:

- Lower secondary, upper secondary and post-secondary non-tertiary education (ED34),
- Tertiary education (ED58).

We construct a composite index that can be interpreted as an approximate average duration of post-secondary studies, based only on ED34 and ED58, defined as:

$$\kappa_{t,g}^{(edu)} = (ED34_{t,g} \times 4 + ED58_{t,g} \times 8)/100\,.$$

The marginal effect of the educational level on mortality is denoted by $\beta_x^{(edu)}$. The next economic factor, $\kappa_{t,g}^{(emp)}$, is the regional employment rate in percentage of the total population aged between 25 and 64 years old. The corresponding marginal effect on mortality is denoted by $\beta_x^{(emp)}$.

We also include in our analysis the regional number of physicians per one hundred thousand inhabitants. The corresponding factor and its marginal effect on mortality are denoted by $\kappa_t^{(phy)}$

and $\beta_x^{(phy)}$, respectively.

The next three variables concern environmental conditions: two related to climate and one to air pollution. We include in the model two factors associated with heating degree days (HDD) and cooling degree days (CDD) over one year. HDD measures the demand for energy required to heat a building, whereas CDD quantifies the energy demand for air conditioning. HDD is calculated by subtracting the daily mean temperature from 18 °C; HDD values are summed over time to estimate heating demand, with higher numbers corresponding to colder regions. Similarly, CDD values are obtained by summing the positive differences between the mean daily temperature and 21 °C, with higher numbers for hotter regions. The HDD and CDD indices are available from Eurostat for each NUTS2 region and are denoted by $HDD_{t,g}$ and $CDD_{t,g}$. To capture differences in regional climates, we consider the following relative ratios:

$$\begin{aligned} \kappa_{t,g}^{(HDD)} &= \frac{HDD_{t,g}}{\frac{1}{g_M}\sum_{g=0}^{g_M} HDD_{t,g}}, \\ \kappa_{t,g}^{(CDD)} &= \frac{CDD_{t,g}}{\frac{1}{g_M}\sum_{g=0}^{g_M} CDD_{t,g}}, \end{aligned}$$

and denote by $\beta_x^{(HDD)}$ and $\beta_x^{(CDD)}$, their marginal impacts on mortality rates. The last environmental factor is the regional emission of NO2, expressed in kilotonnes of CO_2 equivalent per square kilometer. Annual emission volumes are provided by the Emissions Database for Global Atmospheric Research (EDGAR)[3]. NO2 (Nitrogen Dioxide) emissions primarily result from the combustion of fossil fuels and can cause respiratory irritation, leading to symptoms such as coughing and shortness of breath. The surface area of NUTS2 regions in square kilometers is computed using the Python library *geopandas*. We denote the emission factor by $\kappa_t^{(NO2)}$ and its marginal effect on mortality by $\beta_x^{(NO2)}$.

3 B-splines

We need to estimate several curves (α_x, $\beta_{x,g}^{(end)}$, $\kappa_t^{(end)}$, $\beta_x^{(f)}$ for $f \in \mathcal{F}$). Due to the large number of degrees of freedom, estimating a non-parametric version of this model would lead to overfitting. This motivate us to adopt a parametric approach based on B-splines. We first divide the interval between x_m and x_M into m equidistant intervals and denote by d the degree of B splines. There are $m+1$ knots in this interval and if $\Delta_m := \frac{x_M - x_m}{m}$ then the j^{th} knot, x_j, for $j = 0$ up to $m + 2d$ is defined as

$$x_j = x_m - (d - j)\,\Delta_m\,.$$

$B_{j,d}(x)$ is the B-spline or order d associated to the j^{th} knot. Let us recall that B splines are defined by the recurrence:

$$B_{j,0}(x) = \begin{cases} 1 & x_j \le x < x_j \\ 0 & \text{otherwise}\ , \end{cases}$$

and

$$B_{j,n}(x) = \frac{x - x_j}{x_{j+n} - x_j} B_{j,n-1}(x) + \frac{x_{j+n+1} - x}{x_{j+n+1} - x_{j+1}} B_{j+1,n-1}(x)\,.$$

Using B-splines instead of other interpolation functions offers several advantages. B-splines are defined on local support and use piecewise low-degree polynomials combined smoothly, which

[3] edgar.jrc.ec.europa.eu/dataset_ghg2025_nuts2

prevents oscillations and improves numerical stability. We assume that baseline death rates and marginal effects are expressed as linear combinations of B-splines:

$$\begin{cases} \alpha_x = \sum_{j=0}^{m+d-1} c_j^{(\alpha)} B_{j,d}(x)\,, \\ \beta_{x,g}^{(end)} = \sum_{j=0}^{m+d-1} c_{j,g}^{(end)} B_{j,d}(x) & \forall g \in \{1,...,g_M\}, \\ \beta_x^{(f)} = \sum_{j=0}^{m+d-1} c_j^{(f)} B_{j,d}(x), & \forall f \in \mathcal{F}, \end{cases}$$

where $c_j^{(\alpha)}$, $c_j^{(end)}$ and $c_j^{(f)}$ are real coefficients, $\forall f \in \mathcal{F}$ and $j = 0, ...m+d-1$. We denote by Θ, the vector of parameters. This vector is estimated by log-likelihood maximization:

$$\hat{\Theta} \;=\; \arg\max_{\Theta} \ln \mathcal{L}(\Theta)\,,$$

where the Poisson log-likelihood $\mathcal{L}(\Theta)$ is given by a triple sum over ages, calendar years and regions:

$$\begin{aligned} \mathcal{L}(\Theta) &= \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \ln\left(\frac{\exp\left(-\mu_{x,t,g}\, E_{x,t,g}\right) \left(\mu_{x,t,g}\, E_{x,t,g}\right)^{d_{x,t,g}}}{d_{x,t,g}!} \right) \qquad (2) \\ &= \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \left(d_{x,t,g} \ln\left(\mu_{x,t,g}\right) - \mu_{x,t,g}\, E_{x,t,g} + d_{x,t,g} \ln\left(E_{x,t,g}\right) - \log\left(d_{x,t,g}!\right) \right) \\ &\propto \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \left(d_{x,t,g} \ln\left(\mu_{x,t,g}\right) - E_{x,t,g} \exp\left(\ln\left(\mu_{x,t,g}\right)\right) \right)\,. \end{aligned}$$

If we denote by $\hat{d}_{x,t,g} = E_{x,t,g}\mu_{x,t,g}$, the expected number of deaths forecast by the model, this log-likelihood is then proportional to the following term:

$$\mathcal{L}(\Theta) \;\propto\; \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \left(d_{x,t,g} \ln\left(\mu_{x,t,g}\right) - \hat{d}_{x,t,g} \right)\,.$$

We optimize the log-likelihood numerically by canceling the derivative of $\ln \mathcal{L}(\Theta)$ with respect to Θ, with an (approximated[4]) Newton-Raphson algorithm. During the k^{th} iteration, parameter estimates are updated as follows:

$$\Theta^{(k+1)} \;=\; \Theta^{(k)} - \eta(k) \frac{\frac{\partial}{\partial \Theta} \ln \mathcal{L}(\Theta)\big|_{\Theta=\Theta^{(k)}}}{\frac{\partial^2}{\partial \Theta^2} \mathcal{L}(\Theta)\Big|_{\Theta=\Theta^{(k)}}}\,, \qquad (3)$$

where $\eta(k) \in (0,1)$ is an adaptive learning rate with $\eta(0) = \eta_0$ and

$$\eta(k+1) = \begin{cases} \frac{\eta(k)}{2} & \text{if } \mathcal{L}(\Theta^{(k)}) \geq \mathcal{L}(\Theta^{(k+1)})\,, \\ \min\left(\eta(k)\, 1.05\,,\, 2\right) & \text{if } \mathcal{L}(\Theta^{(k)}) < \mathcal{L}(\Theta^{(k+1)})\,. \end{cases}$$

We expand this update rule for each parameter. Let $\widehat{d}_{x,t,g}^{(k)}$ denote the predicted number of deaths after k iterations:

$$\widehat{d}_{x,t,g}^{(k)} = E_{x,t,g} \exp\left(\hat{\alpha}_x^{(k)} + \hat{\beta}_{x,g}^{(end)(k)} \hat{\kappa}_t^{(end)(k)} + \sum_{f \in \mathcal{F}} \hat{\beta}_x^{(f)(k)} \kappa_{t,g}^{(f)} \right)\,.$$

[4]We do not compute the full Hessian. We just evaluate its diagonal.

Using the partial derivatives of the Poisson log-likelihood, we update sequentially the B-spline coefficients for all $f \in \mathcal{F}$ and $j = 0, ...m + d - 1$ as follows:

$$c_j^{(\alpha)(k+1)} = c_j^{(\alpha)(k)} + \eta(k) \frac{\sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \left(d_{x,t,g} - \hat{d}_{x,t,g}^{(k)} \right) B_{j,d}(x)}{\sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \left(B_{j,d}(x) \right)^2 \hat{d}_{x,t,g}^{(k)}},$$

$$c_{j,g}^{(end)(k+1)} = c_{j,g}^{(end)(k)} + \eta(k) \frac{\sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \hat{\kappa}_t^{(end)(k)} \left(d_{x,t,g} - \hat{d}_{x,t,g}^{(k)} \right) B_{j,d}(x)}{\sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \left(\hat{\kappa}_t^{(end)(k)} B_{j,d}(x) \right)^2 \hat{d}_{x,t,g}^{(k)}},$$

$$c_j^{(f)(k+1)} = c_j^{(f)(k)} + \eta(k) \frac{\sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \kappa_t^{(f)} \left(d_{x,t,g} - \hat{d}_{x,t,g}^{(k)} \right) B_{j,d}(x)}{\sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \sum_{g=1}^{g_M} \left(\kappa_t^{(f)} B_{j,d}(x) \right)^2 \hat{d}_{x,t,g}^{(k)}},$$

$$\hat{\kappa}_t^{(end)(k+1)} = \hat{\kappa}_t^{(end)(k)} + \eta(k) \frac{\sum_{x=x_m}^{x_M} \sum_{g=1}^{g_M} \hat{\beta}_{x,g}^{(end)(k)} \left(d_{x,t,g} - \hat{d}_{x,t,g}^{(k)} \right)}{\sum_{x=x_m}^{x_M} \sum_{g=1}^{g_M} \left(\hat{\beta}_{x,g}^{(end)} \right)^2 \hat{d}_{x,t,g}^{(k)}}.$$

The procedure is initialized by setting $c_j^{(\alpha)(0)}$, $c_{j,g}^{(end)(0)}$ and $\kappa_t^{(end)(0)}$ to parameters obtained from an initial quick run of the calibration procedure for a model without economic and environmental factors (see Equation (1) in Section 4. The initial coefficients of the corresponding B-splines, $c_j^{(f)(0)}$, are set to zero.

The algorithm stops when the absolute increase in the log-likelihood becomes smaller than a predetermined tolerance (set to 10^{-2} in illustrations). To enable comparison after completing the estimation procedure, we scale the estimated coefficients $\hat{c}_{j,g}^{(end)}$ so that the sum over ages of the average of $\hat{\beta}_{x,g}^{(end)}$'s across all regions equals 1:

$$\tilde{c}_{j,g}^{(end)} = \frac{\hat{c}_{j,g}^{(end)}}{\sum_{x=x_m}^{x_M} \frac{1}{g_M} \sum_{g=1}^{g_M} \hat{\beta}_{x,g}^{(end)}}.$$

The endogenous factor $\hat{\kappa}_t^{(end)}$ is then adjusted to compensate for this scaling:

$$\tilde{\kappa}_t^{(end)} \quad = \quad \hat{\kappa}_t^{(end)} \left(\sum_{x=x_m}^{x_M} \frac{1}{g_M} \sum_{g=1}^{g_M} \hat{\beta}_{x,g}^{(end)} \right).$$

There are no additional identification constraints. Unlike the original Lee–Carter model, we cannot adapt the coefficients $\hat{c}_j^{(\alpha)}$ to impose that $\sum_{t=t_m}^{t_M} \hat{\kappa}_t^{(end)} = 1$. Because α_x is common to all regions whereas $\hat{\kappa}_t^{(end)}$ interacts with region-specific endogenous effects $\hat{\beta}_{x,g}^{(end)}$.

Let us mention that the proposed framework can easily be extended to P-splines, which are B-splines with a penalty on large variations between successive coefficients:

$$\hat{\Theta} = \arg\max_{\Theta} \left(\ln \mathcal{L}(\Theta) - \lambda \mathcal{P}(\Theta) \right),$$

where $\lambda > 0$ is an hyperparameter and

$$\mathcal{P}(\Theta) = \sum_{f \in \{\alpha, \mathcal{F}\}} \sum_{j=1}^{m+d-1} \left(c_j^{(f)} - c_{j-1}^{(f)} \right)^2 + \sum_{g=1}^{g_M} \sum_{j=1}^{m+d-1} \left(c_{j,g}^{(end)} - c_{j-1,g}^{(end)} \right)^2.$$

In this article, we have instead chosen a model based on B-splines with a limited number of knots to avoid overfitting. This point will be discussed in the numerical illustrations.

4 Numerical illustration

We fit three models to regional mortality rates for continental France (FR), Italy (IT), and Belgium–Netherlands (BE–NL). The latter two countries are grouped together to ensure a sufficiently large population and geographic coverage. FR, IT, and BE–NL comprise 22, 21, and 23 NUTS2 regions, respectively. Due to data availability in Eurostat, we use time windows 2000–2022 for FR, 2000–2023 for IT, and 2001–2022 for BE–NL. Mortality data cover ages from 0 to 82 years. We consider three benchmark models. The first is the Lee–Carter model (LC), estimated on national data. In this model, the log-mortality rate follows the dynamic:

$$\ln\left(\mu_{x,t,g}^{LC}\right) = \alpha_x^{LC} + \beta_x^{LC}\kappa_t^{LC}\,.$$

In the numerical illustrations, α_x, β_x and κ_t are estimated by maximization the sum of Poisson log-likelihoods across regions. This model is non-parametric and therefore involves $2(x_M - x_m) + (t_M - t_m)$ degrees of freedom. To ensure identifiability, the parameters are scaled such that:

$$\sum_{t=t_m}^{t_M} k_t^{LC} = 0 \quad , \quad \sum_{x=x_m}^{x_M} \beta_x^{LC} = 1\,. \tag{4}$$

The second benchmark is similar to model (1) but excludes economic and environmental factors:

$$\ln\left(\mu_{x,t,g}^{LCp}\right) = \alpha_x^{LCp} + \beta_{x,g}^{LCp}\kappa_t^{LCp}\,. \tag{5}$$

In this multi-population parametric Lee-Carter (LCp) model, the curves α_x^{LCp} and $\beta_{x,g}^{LCp}$ are represented using B-splines. This model also serves as the initialization step for the calibration procedure of the Eco-Env LC model. The third benchmark is a variant of the Li and Lee (LL) model. In this multi-population framework, the force of mortality is modeled as:

$$\ln\left(\mu_{x,t,g}^{LL}\right) \quad = \quad \alpha_x^{LL} + \beta_x^{LL}\kappa_t^{LL} + \beta_{x,g}^{LL}\kappa_{g,t}^{LL}\,.$$

As in the LC model, parameters are estimated by iterative maximization of the Poisson log-likelihood, with identifiability constraints applied to ensure proper scaling. Since this model is non-parametric, the number of degrees of freedom is very large: $(g_M + 2)\,(x_M - x_m) + (g_M + 1)\,(t_M - t_m)$. Because the population size in NUTS2 regions is relatively small compared to national populations, this model tends to overfit the data and produces irregular mortality curves. We will revisit this point in the numerical illustrations.

4.1 Goodness of fit

We focus on a model with a small number of knots, setting m to 5, and the degree of B-splines to $d = 3$ to avoid overfitting. In theory, m and d could be selected by optimizing the AIC or BIC but this typically leads to higher values of m, which cause small oscillations in $\beta_x^{(f)}$ for $f \in \mathcal{F}$, a clear signal of overfitting. This point is further discussed in Appendix A. Table 2 reports the goodness-of-fit statistics. For the three geographic areas, including economic and environmental factors results in a significant increase in the log-likelihood compared to the LC and LCp models. We do not expect the Eco-Env model to achieve a higher log-likelihood than the LL model, given its smaller number of parameters. Nevertheless, the parsimony of the Eco-Env LC (264 to 278 parameters) makes it preferable to the LL model (2,437 to 2,603 parameters) based on the Bayesian Information Criterion (BIC).

		N	m	d	dofs	log-likelihood	AIC	BIC
FR	LC	41998	NA	NA	189	-192 912.81	386 203.61	387 837.59
	LCp	41998	5	3	207	-166 834.12	334 082.23	335 871.83
	LL	41998	NA	NA	2521	-159 835.32	324 712.64	346 507.63
	Eco-Env LC	41998	5	3	271	-163 290.36	327 122.73	329 465.62
IT	LC	41832	NA	NA	190	-197 253.77	394 887.54	396 529.41
	LCp	41832	5	3	200	-152 337.57	305 075.15	306 803.43
	LL	41832	NA	NA	2437	-143 624.91	292 123.82	313 182.96
	Eco-Env LC	41832	5	3	264	-150 749.55	302 027.11	304 308.44
	LC	41998	NA	NA	188	-157 807.14	315 990.28	317 615.61
	LCp	41998	5	3	214	-134 666.14	269 760.27	271 610.38
BE NL	LL	41998	NA	NA	2603	-127 711.14	260 628.28	283 132.2
	Eco-Env LC	41998	5	3	278	-133 392.53	267 341.07	269 744.48

Table 2: Statistics of goodness of fit. N: sample size. dofs: number of degrees of freedom. AIC and BIC: Akaike and Bayesian information criteria.

Table 8 reports the mean absolute errors between observed and predicted death rates, weighted by exposures. The weighted mean absolute errors are computed by region as

$$WMAE_g = \frac{1}{\sum_{x,t} E_{x,t,g}} \sum_{x,t} E_{x,t,g} \left| \mu_{x,t,g} - \hat{\mu}_{x,t,g} \right| \ , \ g = 1, ..., G$$

Figure 1 compares the weighted mean absolute errors for French regions obtained with the LC, Eco-Env LC, and LL models. The detailed figures are provided in 8 of Appendix B. Including economic and environmental variables improves the accuracy of the LC model for many regions, such as Nord-Pas de Calais (-70.50% WMAE), Picardie (-54.68%), and Lorraine (-50.0%). WMAEs for the Eco-Env LC model are slightly higher than those of the LL model but remain close in absolute terms. Errors in Limousin and Corse are not reduced by either the Eco-Env LC or LL models. This is due to the volatility of observed death rates, explained by the relatively small population sizes. Corse, with approximately 360,000 inhabitants, and Limousin, with around 710,000 inhabitants, are among the least populous French regions, which increases measurement errors in death rates.

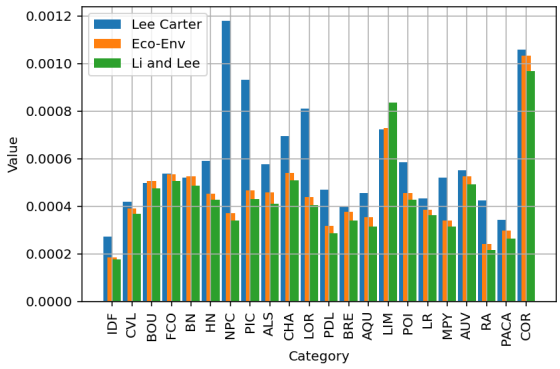


Figure 1: Histogram of weighted mean absolute errors by regions, France.

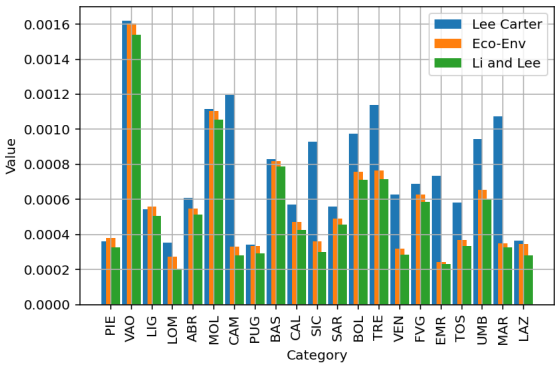


Figure 2: Histogram of weighted mean absolute errors by regions, Italy.

Figure 2 shows regional WMAEs for Italian mortality rates. The conclusions are similar to those for France. We observe an overall decrease in errors with the Eco-Env LC model compared to the LC model. The most significant reductions in WMAE occur in Campania (-73.83%), Marche (-68.02%), Emilia-Romagna (-67.58%), and Sicilia (-64.49%). In Italy, Lombardia is the most populous region (over 10 million inhabitants), while Valle d'Aosta is the least populous (around 115,000 inhabitants). Due to the small size of this region, observed mortality rates exhibit high volatility, which the LC, LL, and Eco-Env models do not replicate. Detailed figures are available in Table 9 in Appendix B. Across all regions, the Eco-Env LC and LL models produce similar error patterns, which is quite remarkable given the limited number of parameters in the Eco-Env LC model (278 degrees of freedom) compared to the LL model (2,603 degrees of freedom).

Figure 3 shows regional WMAEs for Belgium and the Netherlands. These two countries have different economic and health policies; therefore, some economic indicators, such as relative GDP or PPS compared to their averages, may be less relevant than for a single country. Nevertheless, the Eco-Env LC model still considerably reduces the gaps between observed and modeled death rates for many regions. In Belgium, the WMAEs for Hainaut, Liège, and West-Vlaanderen decrease by -54.25%, -47.20%, and -44.96%, respectively. In the Netherlands, the WMAEs for Zeeland, Utrecht, and Noord-Holland fall by -24.81%, -23.54%, and -21.53%. As in France and Italy, the largest errors occur in the least populous regions. In Belgium, these include Luxembourg, Brabant Wallon, and Namur, with approximately 294,000, 413,000, and 502,000 inhabitants, respectively. In the Netherlands, the least populous areas are Zeeland and Flevoland, with 391,000 and 445,000 inhabitants. We again observe similar accuracy for the LL and Eco-Env LC models. Additional figures are available in Table 10 in Appendix B.

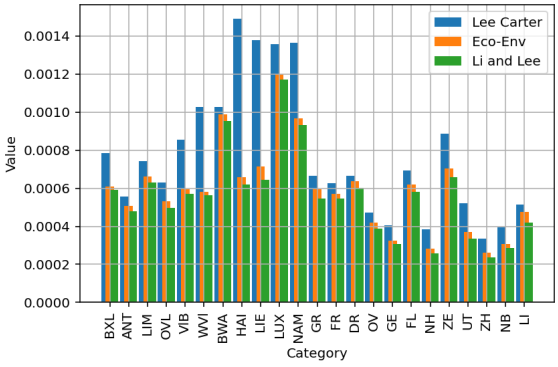


Figure 3: Histogram of weighted mean absolute errors by regions, Belgium and Netherlands.

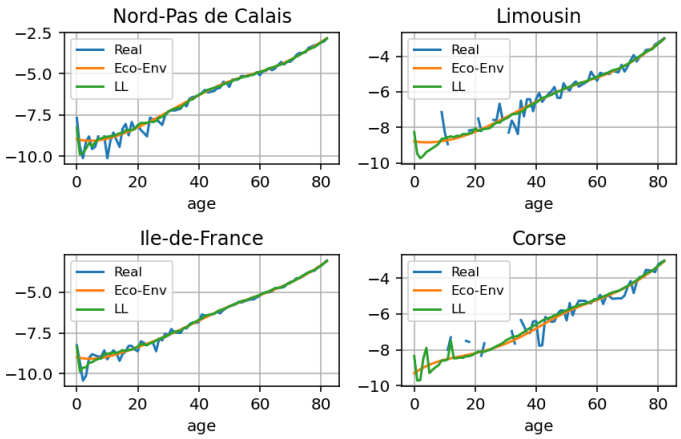


Figure 4: Comparison of Eco-Env LC, LL and observed log-mortality rates for the year 2015, four French regions.

Figure 4 compares modeled and crude log-mortality rates for four French regions in 2015. The Eco-Env LC model produces smooth curves across all ages, a feature that is highly desirable for actuarial applications such as pricing or reserving. In contrast, the LL model generates more erratic mortality curves, particularly at younger ages, which indicates that this non-parametric model slightly overfits the data. As mentioned earlier, Limousin and Corse are the least populous French regions. The number of deaths recorded up to age 40 is therefore small and even zero at some ages, resulting in highly volatile crude rates. This explains why modeling errors in Figure 1 are larger for these regions.

Figures 5, 6 and 7 show maps of WMAEs for the LC and Eco-Env LC models across the three geographic areas considered. These error maps indicate that the model does not favor fitting any specific geographical area but instead achieves an overall, global fit. The WMAEs do not exhibit any particular geographic pattern, and the largest errors are observed in less populous regions. This is particularly evident in southern Belgium, which has fewer inhabitants than neighboring regions.

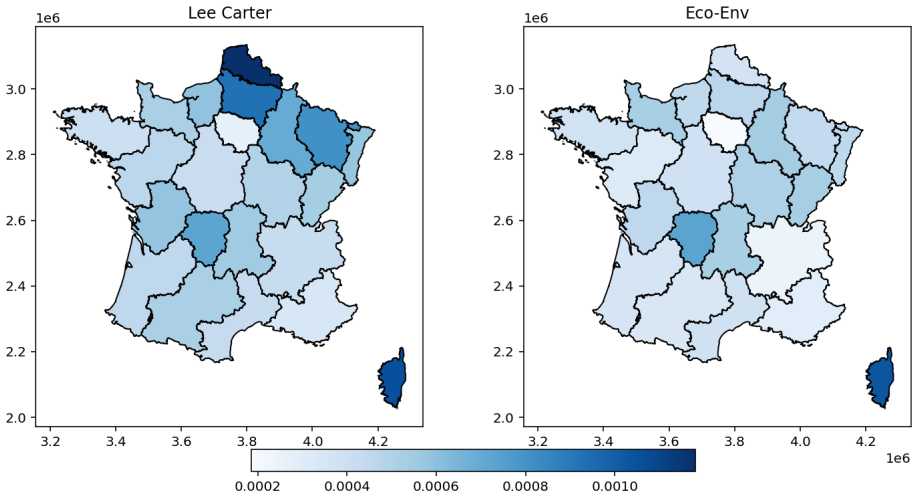


Figure 5: Left and right plots: maps of WMAEs with the Lee-Carter and the Eco-Env Lee Carter models. France.

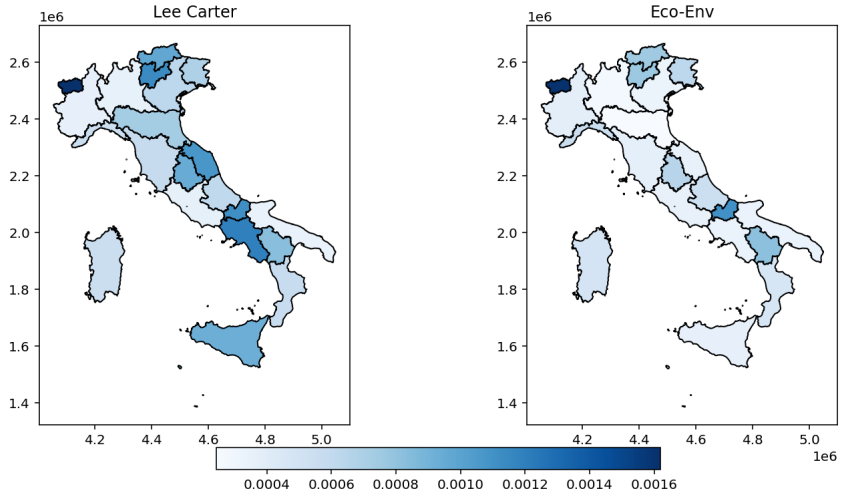


Figure 6: Left and right plots: maps of WMAEs with the Lee-Carter and the Eco-Env Lee Carter models. Italy

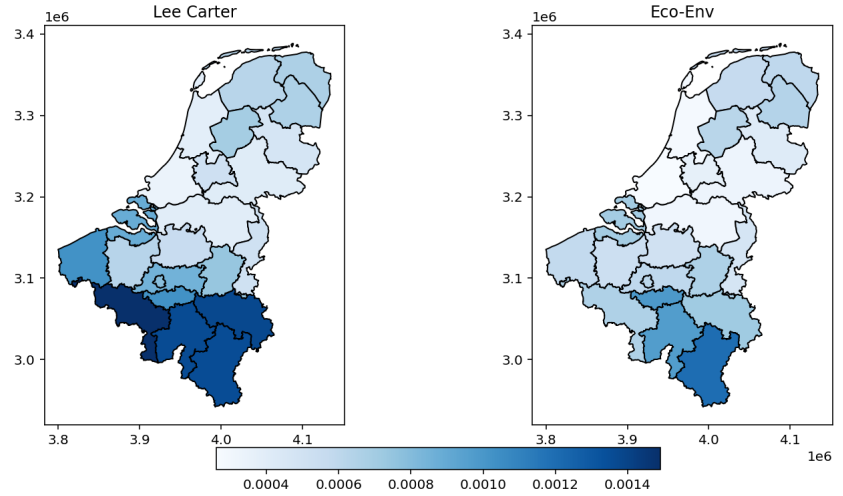


Figure 7: Left and right plots: maps of WMAEs with the Lee-Carter and the Eco-Env Lee Carter models. Italy

4.2 Marginal effects of factors

The left plot of Figure 10 compares the baseline curves, α_x of log-mortality rates for France, Italy, and Belgium–Netherlands. Below age 50, the patterns differ slightly. The right plot of the same figure shows the evolution of the endogenous factor $\kappa_t^{(end)}$, common to all NUTS 2 regions. The curves may have different intercepts because there is no rescaling of $\kappa_t^{(end)}$ (to avoid introducing a region-specific baseline curve $\alpha_{x,g}$). Nevertheless, they exhibit similar decreasing trends, with a parallel shift for BE–NL. The shift is explained by the shorter life expectancy in BE-NL compared to France and Italy. This point is further examined in sub-section 4.3. The peaks of $\kappa_t^{(end)}$ in 2003 for France and Italy correspond to the excess mortality caused by a severe heatwave in August. The bumps in 2020–2021 are attributable to the COVID-19 pandemic.

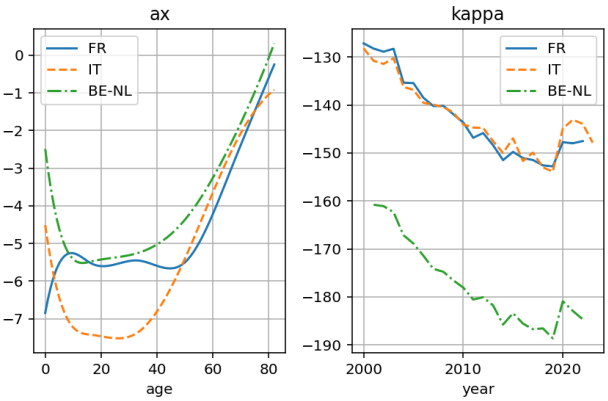


Figure 8: Left and right plots: curves of $\alpha_x^{(end)}$ and $\kappa_t^{(end)}$ for France, Italy, Belgium-Netherlands.

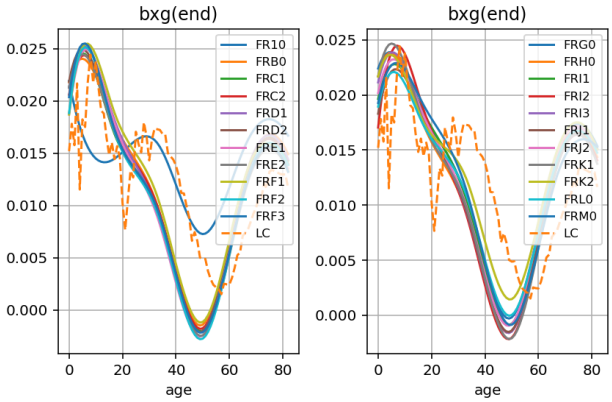


Figure 9: Left and right plots: curves of $\beta_{x,g}^{(end)}$ for French regions.

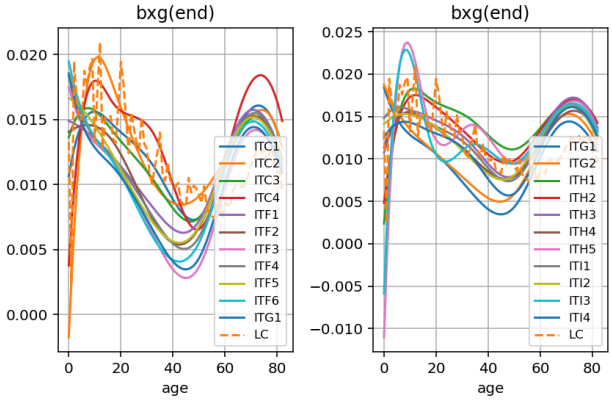


Figure 10: Left and right plots: curves of $\beta_{x,g}^{(end)}$ for Italian regions.

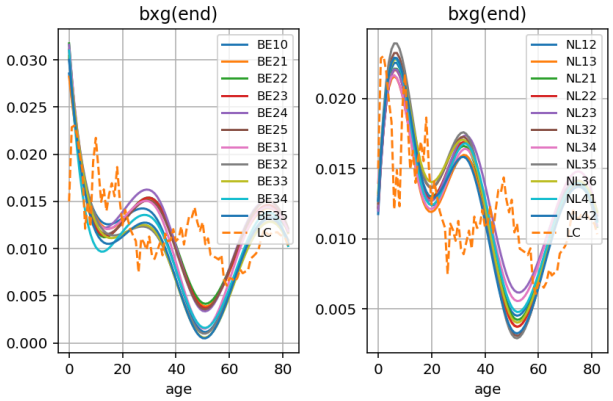


Figure 11: Left and right plots: curves of $\beta_{x,g}^{(end)}$ for Belgian and Netherlands regions.

Figure 9, 10 and 11 show the curves of $\beta_{x,g}^{(end)}$, which represent the marginal effects of $\kappa_t^{(end)}$ on mortality, by region. These are compared to the corresponding curve β_x^{LC} from the Lee-Carter model (legend: LC). Since the Lee–Carter approach is non-parametric, the β_x^{LC}'s display a non-smooth pattern, which is not suitable for actuarial applications. In contrast, we observe a set of smooth, nearly parallel curves for $\beta_{x,g}^{(end)}$. Depending on the country, the overall shape of $\beta_{x,g}^{(end)}$ differs slightly, but all display a decreasing trend between ages 10 and 50, followed by an increase up to age 80.

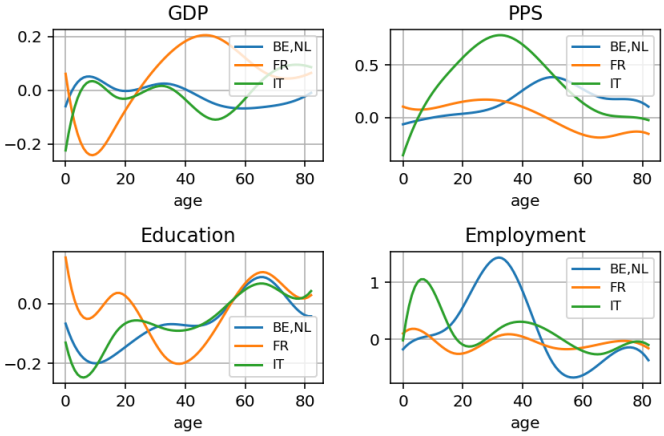


Figure 12: Marginal impact on mortality rates of GDP, PPS, Educational indices and employment rates.

Figure 12 and 13 compare the marginal effects of economic and environmental factors on mortality. Most of these factors exhibit different shapes depending on the country. The ratio of regional to average GDP, $\kappa_t^{(gdp)}$, has a greater impact in France than in the other countries. French regions with $\kappa_t^{(gdp)} > 1$ experience slightly lower mortality up to around age 25 and higher mortality at older ages. In Italy and Belgium-Netherlands, regions with $\kappa_t^{(gdp)} > 1$ have marginally lower mortality rates between ages 40 and 60. This interpretation should be treated with caution, as the GDP ratio is correlated with other economic factors that exert an opposite effect on mortality rates, as emphasized in Rhum (2005).

The ratio of regional over national purchasing power standards, $\kappa_t^{(pps)}$, generally has the same increasing effect on mortality rates across different periods, though with varying amplitudes depending on the country. This is an interesting and counterintuitive observation. However,

GDP and PPS in wealthier regions are correlated with other factors that reduce death rates, such as educational attainment. Indeed, we observe that a high educational index, $\kappa_{t,g}^{(edu)}$, primarily lowers mortality rates, particularly between ages 30 and 55, regardless of the country. The regional employment rate appears less influential in France than in Belgium–Netherlands. In this latter area, a high employment rate marginally increases death rates between ages 5 and 48.

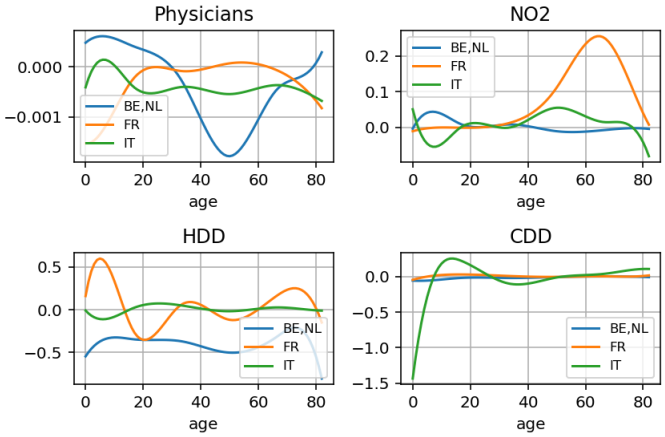


Figure 13: Marginal impact on mortality rates of the number of Physicians per 100 000 inhabitants, NO2 emissions, HDD and CDD.

Figure 13 shows the marginal effect of the number of physicians per 100 000 inhabitants. As expected, a higher number of physicians per capita globally reduces mortality rates. The curves representing the marginal impact of NO_2 pollution are mostly positive, confirming the harmful effect of this gas on health. The ratio of regional to average Heating Degree Days (HDD) does not appear significant for Italy, unlike the ratio of Cooling Degree Days (CDD). This can be explained by the relatively small HDD values compared to CDD in Italy, which has a Mediterranean climate. Conversely, CDD values do not seem very significant for France and Belgium–Netherlands compared to Italy. This likely reflects the low CDD values in these regions, which on average have a colder climate than Italy.

	France		Italy		BE-NL	
	Log. like.	P-values	Log. like.	P-values	Log. like.	P-values
GDP	-163 434	0.0000	-150 827	0.0000	-133 571	0.0000
PPS	-163 338	0.0000	-150 932	0.0000	-133 526	0.0000
Edu.	-164 027	0.0000	-150 936	0.0000	-133 535	0.0000
Emp.	-163 293	0.6629	-150 901	0.0000	-133 407	0.0002
Phys.	-163 384	0.0000	-150 978	0.0000	-133 592	0.0000
NO2	-163 391	0.0000	-150 796	0.0000	-133 415	0.0000
HDD	-163 332	0.0000	-150 776	0.0000	-133 425	0.0000
CDD	-163 307	0.0001	-150 777	0.0000	-133 401	0.0202
Full	-163 290	0.000	-150 750	0.0000	-133 392	0.0000

Table 3: Likelihood ratio test.

To conclude this section, we use a log-likelihood ratio test to assess whether the considered economic and environmental factors are statistically significant. We compare the model including all risk factors with models from which a single risk factor is removed. The full and restricted Eco-Env LC models contain 271 and 263 parameters, respectively. The log-likelihood ratio statistic is then benchmarked against the χ^2-distribution with 8 degrees of freedom. P-values are reported

in Table 3. For France, the employment rate is not a significant factor, apparently because it is redundant with GDP and PPS. For Italy, all factors are highly significant, as p-values are zero. For Belgium and the Netherlands, CDD is not significant at the 1% confidence level. This can be explained by the low number of days in this region with temperatures above 24°C, making CDD less relevant for differentiating climates between regions in BE–NL.

4.3 Evolution of life expectancy

We forecast mortality rates to 2050 using the Eco-Env LC model and compute period life expectancy at birth. Our approach relies on linear regressions rather than simulations. The rationale for this choice is the large number of factors that would need to be simulated. For France, for instance, there are 198 time series of endogenous, economic, and environmental factors. Building a consistent econometric model is challenging, and simulations are computationally time-consuming.

To evaluate the impact of changes in economic and environmental factors, we adopt an incremental approach. First, we linearly extrapolate the endogenous factor $\kappa_t^{(end)}$, while keeping all other factors constant at their last observed values in 2023 or 2022. The coefficient of determination, R^2, for these regressions are equal to 0.86, 0.77 and 0.89 for France, Italy and Belgium-Netherlands, respectively. These high values confirm that a basic linear regression explains the trend of $\kappa_t^{(end)}$ well. The period life expectancies at birth based on these projections are reported in columns e_{50}^{end} of Tables 11, 12 and 10 in Appendix C. The extrapolation is based on $\kappa_t^{(end)}$ values, excluding years 2019 to 2021 to avoid distortions caused by the COVID-19 pandemic. For ages above 82, we linearly extrapolate log-mortality rates up to age 110 using forecast rates for ages between 70 and 82.

Tables 11, 12 and 13 report the maximum and minimum variations, $\Delta_{23/50}^{end}$ or $\Delta_{24/50}^{end}$, in life expectancies between 2023 (e_{23}) or 2024 (e_{24}) and 2050 (e_{50}^{end}). The largest improvements in forecast longevity occur in Ile de France (+5.35 years), Lombardia (+3.07 years), Bruxelles (+4.02 years). The smallest predicted variations are for Picardie (3.53 years), Sicilia (2.32 years) and Luxembourg (3.29 years). We also observe that life expectancy in Belgium–Netherlands in 2023 is generally about two years shorter than in France or Italy.

The Figure 14 presents maps of period life expectancy at birth, for 2023 (e_{23}) and 2050 (e_{50}^{end}).In 2023, the regions with the lowest longevity are located in the northern part of the country: Nord-Pas de Calais (81.18 years), Picardie (81.38 years), and Lorraine (82.00 years). Despite an increase of around three years by 2050, these regions still have the lowest projected life expectancy (84.91, 84.91, and 85.56 years, respectively). Notably, these regions border Belgium, which shows similar figures. The three regions with the highest life expectancies in both 2023 and 2050 are Ile de France (84.69 / 90.04 years), Rhône-Alpes (84.19 / 88.36 years) and Corse (84.04 /88.69 years).

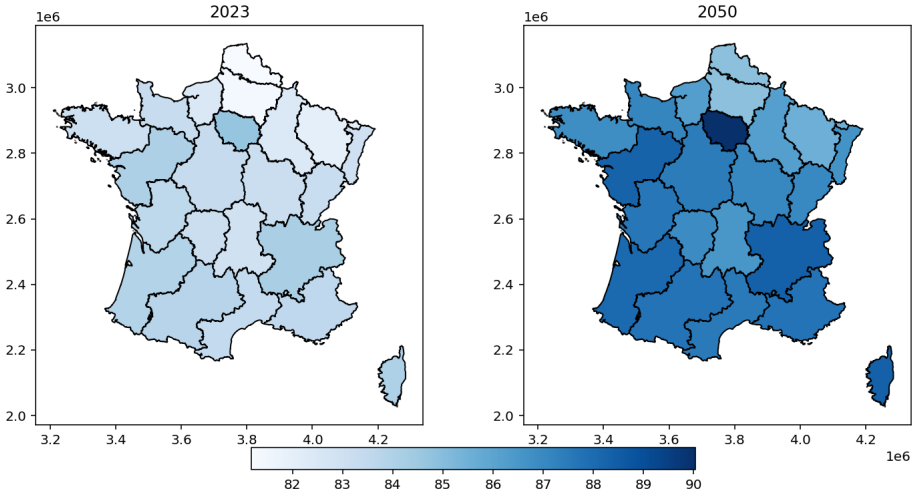


Figure 14: Evolution of period expectancy at birth, from 2023 to 2050, France, extrapolation of $\kappa_t^{(end)}$, only.

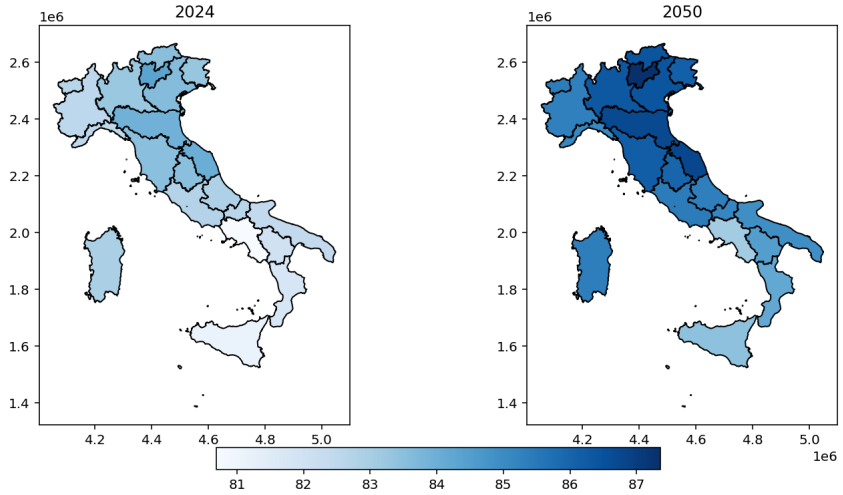


Figure 15: Evolution of period life expectancy at birth, from 2024 to 2050, Italy, extrapolation of $\kappa_t^{(end)}$, only.

Figure 15 compares life expectancy forecasts for 2024 (e_{24}) and 2050 forecasts (e_{50}^{end}), in Italy. Without extrapolating factors other than the endogenous one, the maps reveal that life expectancy remains higher in northern regions, both in 2024 and 2050. In this area, the regions of Trento, Marche, and Emilia-Romagna display the highest longevity in 2024 (84.33, 84.05, and 83.94 years) and in 2050 (87.36, 86.77, and 86.71 years).

Figure 16 highlights a clear divide between southern Belgium, corresponding to the NUTS 1 region called Wallonia, and the northern part of Belgium–Netherlands. In Wallonia, life expectancy is on average two years shorter than in other areas. Despite a significant increase in longevity in southern Belgium, this gap remains visible in 2050. Life expectancies in Wallonia are consistent with those computed using a different model for northern France.

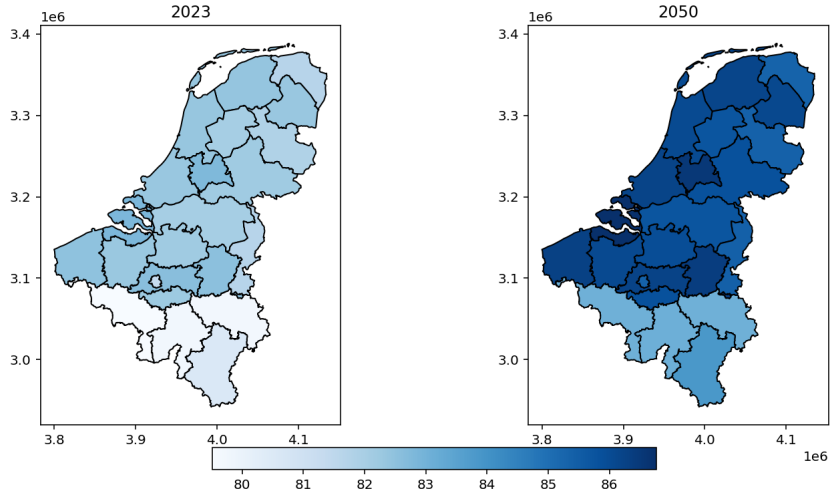


Figure 16: Evolution of period life expectancy at birth, from 2023 to 2050, Belgium-Netherlands, extrapolation of $\kappa_t^{(end)}$, only.

Next, we linearly extrapolate the three environmental factors ($\kappa_t^{(NO2)}$, $\kappa_t^{(HDD)}$, $\kappa_t^{(CDD)}$) and recompute life expectancy at birth, denoted by e_{50}^{env}, for 2050. Tables 14, 15 and 16 report the R^2 of regressions. For $\kappa^{(NO2)}$,the R^2 exceeds 0.70, confirming the predictive capacity of the linear regression. In contrast, the R^2 values for $\kappa_t^{(HDD)}$ and $\kappa_t^{(CDD)}$ are close to zero. These factors are essentially constant over time, so the best prediction is simply the mean of the observed $\kappa_t^{(HDD)}$ and $\kappa_t^{(CDD)}$. This explains why the corresponding R^2 values are nearly null. The upper and lower variations $\Delta_{50/50}^{env} = e_{50}^{env} - e_{50}^{end}$ are presented in Tables 11, 12 and 13. For France, Italy, and Belgium–Netherlands, the largest improvements in life expectancy attributable to environmental factors are 0.32, 0.07, and 0.64 years, respectively. The largest reductions caused by changes in environmental factors are -0.23, -0.80, and -0.69 years. These results are far from negligible because the effects of environmental factors on mortality rates are dampened by the calculation of life expectancy. Indeed, these factors do not affect death rates uniformly across all ages. It is also important to note that these forecasts are extrapolated from data covering the past two decades and therefore do not account for excess mortality related to the acceleration of climate change.

Finally, we linearly extrapolate the economic factors ($\kappa_t^{(GDP)}$, $\kappa_t^{(PPS)}$, $\kappa_t^{(edu)}$,$\kappa_t^{(phys)}$) and recalculate life expectancy at birth, denoted by e_{50}^{eco}, for 2050. For most of factors and countries, the corresponding R^2 ranges from 0.60 to nearly 1.00, which is acceptable. The time series of physicians per 10 000 inhabitants in France display volatile sample paths for a few regions, reducing the efficiency of the linear regression. For some factors and regions, theR^2 is nearly zero because the time series are flat. For these factors, the historical mean is the best predictor of future values. The smaller and larger variations of life expectancies, $\Delta_{50/50}^{eco} = e_{50}^{eco} - e_{50}^{env}$ are shown in Tables 11, 12 and 13. The greatest increases in longevity related to economic factors are predicted for Corse (+1.23 years), Lazio (+1.50 years), and Brabant Wallon (+1.52 years). In Ile-de-France, the evolution of economic conditions is expected to reduce life expectancy by -0.74. In Italy and the Netherlands, the smallest increase in longevity linked to economic evolution occurs in Basilicata (+0.21 years) and Utrecht (+0.25 years).

France				
		2023, e_{23}	2050, e_{50}^{end}	$\Delta_{23/50}^{end}$
Max $\Delta_{23/50}^{end}$	Ile de France	84.69	90.04	5.35
Min $\Delta_{23/50}^{end}$	Picardie	81.38	84.91	3.53
		2050, e_{50}^{end}	2050, e_{50}^{env}	$\Delta_{50/50}^{env}$
Max $\Delta_{50/50}^{env}$	Corse	88.37	88.69	0.32
Min $\Delta_{50/50}^{env}$	Midi-Pyrénées	87.75	87.52	-0.23
		2050, e_{50}^{env}	2050, e_{50}^{eco}	$\Delta_{50/50}^{eco}$
Max $\Delta_{50/50}^{eco}$	Corse	88.69	89.92	1.23
Min $\Delta_{50/50}^{eco}$	Ile de France	90.07	89.33	-0.74

Table 4: Evolution of life expectancy, France.

Italy				
		2024, e_{24}	2050, e_{50}^{end}	$\Delta_{24/50}^{end}$
Max $\Delta_{24/50}^{end}$	Lombardia	83.26	86.33	3.07
Min $\Delta_{24/50}^{end}$	Sicilia	81.15	83.47	2.32
		2050, e_{50}^{end}	2050, e_{50}^{env}	$\Delta_{50/50}^{env}$
Max $\Delta_{50/50}^{env}$	Abruzzo	85.34	85.41	0.07
Min $\Delta_{50/50}^{env}$	Venetor	86.41	85.61	-0.80
		2050, e_{50}^{env}	2050, e_{50}^{eco}	$\Delta_{50/50}^{eco}$
Max $\Delta_{50/50}^{eco}$	Lazio	85.40	86.90	1.50
Min $\Delta_{50/50}^{eco}$	Basilicata	84.46	84.67	0.21

Table 5: Evolution of life expectancy, Italy.

Belgium-Netherlands				
		2023, e_{23}	2050, e_{50}^{end}	$\Delta_{23/50}^{end}$
Max $\Delta_{23/50}^{end}$	Bruxelles	81.34	85.36	4.02
Min $\Delta_{23/50}^{end}$	Luxembourg	80.58	83.87	3.29
		2050, e_{50}^{end}	2050, e_{50}^{env}	$\Delta_{50/50}^{env}$
Max $\Delta_{50/50}^{env}$	Luxembourg	83.87	84.51	0.64
Min $\Delta_{50/50}^{env}$	Noord-Holland	86.06	85.37	-0.69
		2050, e_{50}^{env}	2050, e_{50}^{eco}	$\Delta_{50/50}^{eco}$
Max $\Delta_{50/50}^{eco}$	Brabant Wallon	85.73	87.25	1.52
Min $\Delta_{50/50}^{eco}$	Utrecht	86.51	86.76	0.25

Table 6: Evolution of life expectancy, Belgium-Netherlands.

5 Conclusions

This study introduced a parametric multi-population mortality model that integrates economic and environmental determinants into the Lee–Carter framework. By incorporating eight external risk factors and modeling their marginal effects through B-splines, the proposed Eco-Env LC model achieves a balance between interpretability and parsimony. Compared to non-parametric

approaches such as Li and Lee, our method avoids overfitting and produces smooth mortality curves, which are essential for actuarial and demographic applications.

Empirical results for France, Italy, and Belgium–Netherlands confirm the relevance of economic and environmental variables in explaining regional mortality differences. The inclusion of these factors significantly improves goodness-of-fit metrics relative to traditional LC and parametric LC models, while maintaining a much lower number of parameters than the Li and Lee specification. Weighted mean absolute errors demonstrate that the Eco-Env LC model reduces discrepancies for most regions, particularly those with large populations, and performs comparably to the Li and Lee model despite its simpler structure.

Analysis of marginal effects reveals heterogeneous impacts across countries and age groups. Educational attainment and physician density consistently reduce mortality, while NO2 emissions exert a harmful influence. Climate-related indicators (HDD and CDD) show country-specific relevance, reflecting regional climatic diversity. Economic indicators such as GDP and PPS exhibit complex interactions, sometimes counterintuitive, due to correlations with other socio-economic factors.

Forecasts of life expectancy to 2050 highlight persistent regional disparities, even under optimistic scenarios. While longevity is projected to increase everywhere, improvements remain uneven, with southern Belgium and certain French regions lagging behind. Incremental simulations suggest that environmental trends will have modest but non-negligible effects on future mortality, whereas economic evolution could amplify regional gaps if inequalities persist.

Overall, this work demonstrates that integrating exogenous determinants into mortality modeling enhances explanatory power and provides a framework for assessing the demographic consequences of economic and environmental policies. Future research could extend the model to include lifestyle factors, refine climate projections under accelerated warming, and explore stochastic formulations to capture uncertainty in long-term forecasts.

Appendix A

In this appendix, we discuss the influence of the number m of knots and the degree d of B-splines, on the goodness-of-fit statistics.

France						
	m	d	dofs	log-likelihood	AIC	BIC
LC	NA	NA	189	-192 912.81	386 203.61	387 837.59
LL	NA	NA	2521	-159 835.32	324 712.64	346 507.63
Eco-Env LC	5	3	271	-163 290.36	327 122.73	329 465.62
Eco-Env LC	10	3	426	-160 197.08	321 246.15	324 929.08
Eco-Env LC	15	3	581	-159 412.94	319 987.89	325 010.85
Eco-Env LC	10	2	395	-161 768.36	324 326.72	327 741.65
Eco-Env LC	10	3	426	-160 197.08	321 246.15	324 929.08
Eco-Env LC	10	4	457	-159 483.94	319 881.88	323 832.82

Table 7: Statistics of goodness of fit. dofs: number of degrees of freedom. AIC and BIC: Akaike and Bayesian information criteria.

Table 7 reports the log-likelihood, AIC, and BIC for the Eco-Env LC model fitted to French regional data. It shows that increasing the number of knots significantly improves the log-

likelihood. With $m = 15$, we even achieve a slightly lower log-likelihood than the LL model. However, in this configuration, the large number of parameters penalizes the BIC, which becomes higher than with 10 knots. This is the first indication that models with large m should not be considered.

This intuition is confirmed by Figures 8, 9 and 11. Increasing the number of fixed knots leads to oscillations in the curves of α_x, $\beta_{x,g}^{(PPS)}$, $\beta_{x,g}^{(emp)}$, $\beta_{x,g}^{(phys)}$, $\beta_{x,g}^{(NO2)}$, and $\beta_{x,g}^{(CDD)}$. This is another clear sign of overfitting. Ideally, the predictive capacity of the model for various values of m should be evaluated on a test set. However, this validation procedure is not feasible given the restricted time window for which data are available. We have therefore based our analysis on a parsimonious model with only five knots. Table 7 also shows that increasing the degree of B-splines has a positive impact on goodness-of-fit statistics. Nevertheless, for $d = 4$, we also observe oscillations in the curves of α_x, $\beta_{x,g}^{(f)}$ for $f \in \mathcal{F}$ and therefore opt for B-splines of degree three.

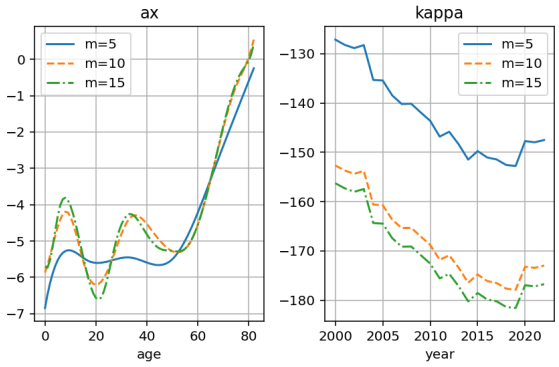


Figure 17: Left and right plots: curves of $\alpha_x^{(end)}$ and $\kappa_t^{(end)}$ for France, Italy, Belgium-Netherlands.

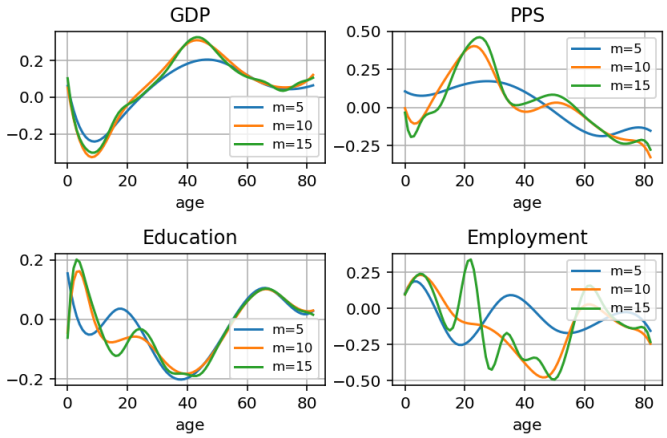


Figure 18: Marginal impact on mortality rates of GDP, PPS, Educational indices and employment rates.

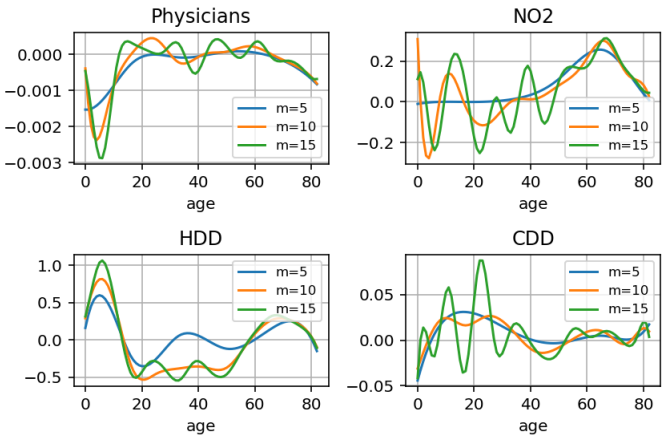


Figure 19: Marginal impact on mortality rates of the number of Physicians per 100 000 inhabitants, NO2 emissions, HDD and CDD.

Appendix B

This section contains tables of mean absolute errors between observed and predicted death rates, weighted by exposures, for France, Italy, and Belgium–Netherlands.

Regions		Lee-Carter	Eco-Env Lee-Carter	Li & Lee	Variation (%)	
					LC to Eco-Env LC	LC to LL
Ile-de-France	IDF	0.0004	0.0003	0.0003	-38.52	-41.83
Centre - Val de Loire	CVL	0.0005	0.0005	0.0004	-8.90	-14.25
Bourgogne	BOU	0.0006	0.0006	0.0005	2.70	-4.18
Franche-Comté	FCO	0.0006	0.0006	0.0006	0.09	-5.41
Basse-Normandie	BN	0.0006	0.0006	0.0006	2.09	-6.10
Haute-Normandie	HN	0.0008	0.0006	0.0005	-24.15	-28.77
Nord-Pas de Calais	NPC	0.0017	0.0005	0.0005	-70.5	-73.11
Picardie	PIC	0.0013	0.0006	0.0005	-54.68	-59.11
Alsace	ALS	0.0008	0.0006	0.0005	-25.07	-36.29
Champagne	CHA	0.0008	0.0006	0.0006	-22.48	-27.05
Lorraine	LOR	0.0011	0.0005	0.0005	-50.00	-54.50
Pays de la Loire	PDL	0.0006	0.0004	0.0004	-34.41	-41.42
Bretagne	BRE	0.0005	0.0004	0.0004	-7.15	-17.43
Aquitaine	AQU	0.0005	0.0004	0.0004	-25.39	-34.3
Limousin	LIM	0.0008	0.0008	0.0009	0.60	17.71
Poitou-Charentes	POI	0.0007	0.0005	0.0005	-23.99	-29.66
Languedoc-Roussillon	LR	0.0005	0.0004	0.0004	-11.41	-16.72
Midi-Pyrénées	MPY	0.0006	0.0004	0.0004	-35.32	-41.17
Auvergne	AUV	0.0006	0.0006	0.0005	-6.33	-13.51
Rhône-Alpes	RA	0.0005	0.0003	0.0003	-40.16	-47.41
Provence-Alpes-Côte d'Azur	PACA	0.0004	0.0003	0.0003	-14.52	-24.63
Corse	COR	0.0012	0.0011	0.0011	-2.27	-9.03
Global		0.0006	0.0005	0.0005	-27.25	-32.17

Table 8: Table of weighted mean absolute errors (WMAE), France.

Regions		Lee-Carter	Eco-Env Lee-Carter	Li & Lee	Variation (%) LC to Eco-Env LC	LC to LL
PIE	Piemonte	0.0004	0.0004	0.0004	5.84	-8.86
VAO	Valle d'Aosta	0.0017	0.0017	0.0016	-1.12	-4.85
LIG	Liguria	0.0005	0.0005	0.0005	2.89	-7.23
LOM	Lombardia	0.0004	0.0003	0.0002	-23.72	-44.98
ABR	Abruzzo	0.0007	0.0006	0.0006	-9.97	-15.17
MOL	Molise	0.0012	0.0012	0.0011	-0.93	-5.48
CAM	Campania	0.0017	0.0005	0.0004	-73.83	-78.45
PUG	Puglia	0.0004	0.0004	0.0003	-2.05	-15.92
BAS	Basilicata	0.0009	0.0009	0.0009	-2.13	-4.73
CAL	Calabria	0.0007	0.0006	0.0005	-21.98	-29.99
SIC	Sicilia	0.0012	0.0004	0.0004	-64.49	-70.92
SAR	Sardegna	0.0007	0.0006	0.0005	-12.59	-18.5
BOL	Pr. di Bolzano	0.0012	0.0009	0.0009	-23.48	-28.8
TRE	Pr. di Trento	0.0014	0.0009	0.0008	-36.05	-41.16
VEN	Veneto	0.0008	0.0004	0.0003	-51.35	-56.21
FVG	Friuli-Venezia Giulia	0.0007	0.0006	0.0006	-10.74	-16.89
EMR	Emilia-Romagna	0.0008	0.0003	0.0003	-67.58	-69.42
TOS	Toscana	0.0006	0.0004	0.0003	-37.86	-43.82
UMB	Umbria	0.001	0.0007	0.0006	-31.68	-37.61
MAR	Marche	0.0012	0.0004	0.0004	-68.02	-70.47
LAZ	Lazio	0.0004	0.0004	0.0003	-7.23	-25.31
	Global	0.0006	0.0006	0.0006	-30.46	-36.47

Table 9: Table of weighted mean absolute errors (WMAE), Italy.

Regions		Lee-Carter	Eco-Env Lee-Carter	Li & Lee	Variation (%) LC to Eco-Env LC	LC to LL
BXL	Bruxelles	0.0011	0.0009	0.0009	-19.13	-20.81
ANT	Antwerpen	0.0007	0.0006	0.0006	-7.01	-13.04
LIM	Limburg	0.0009	0.0008	0.0008	-11.33	-15.69
OVL	Oost-Vlaanderen	0.0007	0.0006	0.0006	-14.38	-21.09
VlB	Vlaams Brabant	0.0011	0.0007	0.0007	-32.99	-36.23
WVl	West-Vlaanderen	0.0012	0.0006	0.0006	-44.96	-47.23
BWA	Brabant wallon	0.0013	0.0012	0.0012	-3.57	-6.87
HAI	Hainaut	0.0018	0.0008	0.0008	-54.25	-56.85
LIE	Liège	0.0017	0.0009	0.0008	-47.2	-53.16
LUX	Luxembourg	0.0017	0.0015	0.0015	-9.24	-11.08
NAM	Namur	0.0016	0.0012	0.0012	-26.79	-29.36
GR	Gronigen	0.0009	0.0008	0.0007	-13.43	-19.08
FR	Friesland	0.0008	0.0007	0.0007	-8.31	-11.27
DR	Drenthe	0.0008	0.0008	0.0007	-5.24	-10.35
OV	Overijssel	0.0006	0.0005	0.0005	-11.03	-17.15
GE	Gelderland	0.0005	0.0004	0.0004	-15.38	-18.92
FL	Flevoland	0.0011	0.0011	0.001	-5.65	-9.83
NH	Noord-Holland	0.0005	0.0004	0.0003	-21.53	-26.44
ZE	Zeeland	0.0011	0.0008	0.0008	-24.81	-30.1
UT	Utrecht	0.0007	0.0005	0.0005	-23.54	-30.21
ZH	Zuid-Holland	0.0004	0.0003	0.0003	-19.45	-24.2
NB	Noord-Brabant	0.0005	0.0004	0.0004	-17.49	-22.49
LI	Limburg	0.0006	0.0006	0.0005	-9.9	-22.1
	Global	0.0008	0.0007	0.0007	-22.25	-26.45

Table 10: Table of weighted mean absolute errors (WMAE), Belgium-Netherlands.

Appendix C

This section presents life expectancy forecasts with and without extrapolation of economic and environmental factors for France, Italy, and Belgium–Netherlands.

	2023	2050					
		Extrapolated factors			Variations		
	e_{23}	e_{50}^{end}	e_{50}^{env}	e_{50}^{eco}	$\Delta_{23/50}^{end}$	$\Delta_{50/50}^{env}$	$\Delta_{50/50}^{eco}$
IDF	84.69	90.04	90.07	89.33	4.64	0.03	-0.74
CVL	83.38	87.47	87.47	87.44	4.06	0	-0.03
BOU	83.19	87.09	87.27	87.46	4.27	0.18	0.19
FCO	83.28	87.06	87.17	87.61	4.33	0.11	0.44
BN	83.33	87.14	87.22	88.09	4.76	0.08	0.87
HN	82.47	86.26	86.46	86.68	4.21	0.2	0.22
NPC	81.18	84.91	85.06	85.61	4.43	0.15	0.55
PIC	81.38	84.91	85.02	85.06	3.68	0.11	0.04
ALS	83.03	86.67	86.66	86.91	3.88	-0.01	0.25
CHA	82.45	86.21	86.27	86.41	3.96	0.06	0.14
LOR	82	85.56	85.56	85.85	3.85	0	0.29
PDL	84.1	88.29	88.2	88.85	4.75	-0.09	0.65
BRE	83.07	86.83	86.84	87.68	4.61	0.01	0.84
AQU	83.93	88.1	87.97	88.7	4.77	-0.13	0.73
LIM	83.22	86.91	86.87	87	3.78	-0.04	0.13
POI	83.62	87.61	87.48	87.87	4.25	-0.13	0.39
LR	83.44	87.54	87.42	88.06	4.62	-0.12	0.64
MPY	83.82	87.75	87.52	87.83	4.01	-0.23	0.31
AUV	82.93	86.52	86.42	86.92	3.99	-0.1	0.5
RA	84.19	88.36	88.35	88.87	4.68	-0.01	0.52
PACA	83.6	87.73	87.93	88.18	4.58	0.2	0.25
COR	84.04	88.37	88.69	89.92	5.88	0.32	1.23

Table 11: Life expectancy at birth, in years, France.

	2024	2050					
		Extrapolated factors			Variations		
	e_{24}	e_{50}^{end}	e_{50}^{env}	e_{50}^{eco}	$\Delta_{24/50}^{end}$	$\Delta_{50/50}^{env}$	$\Delta_{50/50}^{eco}$
PIE	82.55	85.31	85.24	86.22	3.67	-0.07	0.98
VAO	82.67	85.4	85.38	86.31	3.64	-0.02	0.93
LIG	82.53	85.12	85.07	86.21	3.68	-0.05	1.14
LOM	83.26	86.33	86.27	87.13	3.87	-0.06	0.86
ABR	82.87	85.34	85.41	86.36	3.49	0.07	0.95
MOL	82.65	85.15	85.14	85.9	3.25	-0.01	0.76
CAM	80.69	83.11	83.09	84.15	3.46	-0.02	1.06
PUG	82.42	84.92	84.87	85.72	3.3	-0.05	0.85
BAS	82.04	84.52	84.46	84.67	2.63	-0.06	0.21
CAL	81.77	84.19	84.17	84.69	2.92	-0.02	0.52
SIC	81.15	83.47	83.24	84.23	3.08	-0.23	0.99
SAR	82.91	85.37	85.36	86.41	3.5	-0.01	1.05
BOL	83.55	86.42	86.42	86.92	3.37	0	0.5
TRE	84.33	87.36	87.42	88.09	3.76	0.06	0.67
VEN	83.59	86.41	85.61	86.46	2.87	-0.8	0.85
FVG	83.3	86.11	86.05	87.28	3.98	-0.06	1.23
EMR	83.94	86.71	86.49	87.59	3.65	-0.22	1.1
TOS	83.52	86.17	86.13	87.34	3.82	-0.04	1.21
UMB	83.52	86	85.92	86.97	3.45	-0.08	1.05
MAR	84.05	86.77	86.76	87.41	3.36	-0.01	0.65
LAZ	82.74	85.42	85.4	86.9	4.16	-0.02	1.5

Table 12: Life expectancy at birth, in years, Italy.

	2023	2050					
		Extrapolated factors			Variations		
	e_{23}	e_{50}^{end}	e_{50}^{env}	e_{50}^{eco}	$\Delta_{23/50}^{end}$	$\Delta_{50/50}^{env}$	$\Delta_{50/50}^{eco}$
BXL	81.34	85.36	85.48	86.67	5.33	0.12	1.19
ANT	82.18	85.95	85.92	87.22	5.04	-0.03	1.3
LIM	82.57	86.4	86.41	87.75	5.18	0.01	1.34
OVL	82.29	86.08	86.2	87.57	5.28	0.12	1.37
VlB	82.62	86.24	86.35	87.43	4.81	0.11	1.08
WVl	82.51	86.29	86.45	87.48	4.97	0.16	1.03
BWA	82.26	85.86	85.73	87.25	4.99	-0.13	1.52
HAI	79.51	83.07	83.3	84.26	4.75	0.23	0.96
LIE	79.72	83.08	83.21	84.12	4.4	0.13	0.91
LUX	80.58	83.87	84.51	85.37	4.79	0.64	0.86
NAM	79.75	83.1	83.2	84.18	4.43	0.1	0.98
GR	81.69	85.33	84.91	85.9	4.21	-0.42	0.99
FR	82.39	86.2	85.9	87.13	4.74	-0.3	1.23
DR	82.34	86.09	85.86	87.17	4.83	-0.23	1.31
OV	81.83	85.4	85.24	86.46	4.63	-0.16	1.22
GE	82.19	85.85	85.9	86.98	4.79	0.05	1.08
FL	82.04	85.74	85.68	86.46	4.42	-0.06	0.78
NH	82.37	86.06	85.37	86.46	4.09	-0.69	1.09
ZE	82.84	86.77	86.73	87.92	5.08	-0.04	1.19
UT	82.78	86.53	86.51	86.76	3.98	-0.02	0.25
ZH	82.33	86.22	86.24	86.76	4.43	0.02	0.52
NB	82.05	85.73	85.61	86.73	4.68	-0.12	1.12
LI	81.69	85.41	85.36	86.25	4.56	-0.05	0.89

Table 13: Life expectancy at birth, Belgium-Netherlands.

Appendix D

This section presents the R^2 of linear regressions of endogenous and environmental factors for France, Italy, and Belgium–Netherlands.

	$\kappa_t^{(end)}$	$\kappa_t^{(gdp)}$	$\kappa_t^{(pps)}$	$\kappa_t^{(edu)}$	$\kappa_t^{(emp)}$	$\kappa_t^{(phys)}$	$\kappa_t^{(NO2)}$	$\kappa_t^{(HDD)}$	$\kappa_t^{(CDD)}$
IDF	0.86	0.58	0.68	0.61	0.96	0.01	0.75	0.15	0.0
CVL	0.86	0.98	0.9	0.02	0.97	0.0	0.61	0.0	0.02
BOU	0.86	0.94	0.56	0.3	0.94	0.06	0.71	0.18	0.0
FCO	0.86	0.93	0.0	0.14	0.94	0.29	0.76	0.05	0.06
BN	0.86	0.89	0.63	0.19	0.93	0.51	0.76	0.21	0.0
HN	0.86	0.92	0.65	0.01	0.93	0.14	0.82	0.1	0.02
NPC	0.86	0.68	0.6	0.3	0.99	0.29	0.84	0.1	0.02
PIC	0.86	0.93	0.51	0.04	0.92	0.1	0.71	0.16	0.06
ALS	0.86	0.83	0.35	0.04	0.94	0.3	0.81	0.05	0.15
CHA	0.86	0.85	0.28	0.06	0.95	0.04	0.62	0.07	0.12
LOR	0.86	0.97	0.07	0.02	0.92	0.03	0.81	0.01	0.14
PDL	0.86	0.86	0.12	0.57	0.98	0.56	0.8	0.01	0.03
BRE	0.86	0.01	0.41	0.62	0.95	0.51	0.79	0.1	0.02
AQU	0.86	0.68	0.04	0.75	0.96	0.53	0.83	0.07	0.02
LIM	0.86	0.92	0.77	0.28	0.97	0.08	0.89	0.02	0.0
POI	0.86	0.36	0.14	0.25	0.96	0.29	0.65	0.17	0.03
LR	0.86	0.61	0.07	0.64	0.92	0.67	0.41	0.33	0.01
MPY	0.86	0.82	0.4	0.55	0.96	0.32	0.7	0.33	0.0
AUV	0.86	0.78	0.53	0.41	0.95	0.27	0.81	0.21	0.12
RA	0.86	0.83	0.00	0.76	0.94	0.59	0.8	0.13	0.01
PACA	0.86	0.41	0.55	0.87	0.97	0.13	0.62	0.02	0.17
COR	0.86	0.88	0.75	0.77	0.79	0.17	0.91	0.02	0.07

Table 14: R^2 linear regression of endogenous and other factors, France.

	$\kappa_t^{(end)}$	$\kappa_t^{(gdp)}$	$\kappa_t^{(pps)}$	$\kappa_t^{(edu)}$	$\kappa_t^{(emp)}$	$\kappa_t^{(phys)}$	$\kappa_t^{(NO2)}$	$\kappa_t^{(HDD)}$	$\kappa_t^{(CDD)}$
PIE	0.77	0.8	0.22	0.02	0.96	0.76	0.9	0.08	0.09
VAO	0.77	0.16	0.85	0.8	0.98	0.78	0.96	0.06	0.01
LIG	0.77	0.8	0.03	0.85	0.86	0.8	0.36	0.07	0.21
LOM	0.77	0.88	0.55	0.52	0.97	0.69	0.44	0.02	0.0
ABR	0.77	0.03	0.06	0.23	0.92	0.8	0.91	0.1	0.32
MOL	0.77	0.85	0.16	0.01	0.97	0.83	0.94	0.01	0.09
CAM	0.77	0.91	0.69	0.34	0.98	0.83	0.83	0.0	0.03
PUG	0.77	0.74	0.66	0.16	0.98	0.83	0.58	0.06	0.02
BAS	0.77	0.53	0.48	0.44	0.95	0.96	0.9	0.13	0.0
CAL	0.77	0.93	0.28	0.11	0.93	0.92	0.9	0.07	0.1
SIC	0.77	0.81	0.54	0.07	0.97	0.83	0.61	0.03	0.25
SAR	0.77	0.44	0.84	0.42	0.98	0.85	0.83	0.01	0.06
BOL	0.77	0.94	0.82	0.94	0.92	0.73	0.95	0.22	0.0
TRE	0.77	0.71	0.61	0.94	0.96	0.82	0.94	0.24	0.23
VEN	0.77	0.7	0.6	0.8	0.97	0.7	0.67	0.0	0.0
FVG	0.77	0.61	0.11	0.9	0.98	0.74	0.92	0.06	0.02
EMR	0.77	0.84	0.56	0.53	0.69	0.69	0.75	0.08	0.07
TOS	0.77	0.25	0.47	0.6	0.96	0.76	0.93	0.0	0.08
UMB	0.77	0.93	0.86	0.73	0.95	0.81	0.8	0.0	0.12
MAR	0.77	0.61	0.87	0.54	0.68	0.77	0.92	0.0	0.02
LAZ	0.77	0.07	0.47	0.75	0.94	0.73	0.77	0.04	0.02

Table 15: R^2, linear regression of endogenous and other factors, Italy.

	$\kappa_t^{(end)}$	$\kappa_t^{(gdp)}$	$\kappa_t^{(pps)}$	$\kappa_t^{(edu)}$	$\kappa_t^{(emp)}$	$\kappa_t^{(phys)}$	$\kappa_t^{(NO2)}$	$\kappa_t^{(HDD)}$	$\kappa_t^{(CDD)}$
BXL	0.89	0.93	0.88	0.54	0.83	0.97	0.15	0.02	0.11
ANT	0.89	0.84	0.26	0.82	0.95	1.0	0.91	0.04	0.0
LIM	0.89	0.87	0.23	0.87	0.95	1.0	0.3	0.07	0.01
OVL	0.89	0.82	0.24	0.88	0.97	1.0	0.09	0.18	0.0
VlB	0.89	0.78	0.33	0.84	0.94	0.77	0.57	0.03	0.15
WVl	0.89	0.84	0.45	0.86	0.99	1.0	0.42	0.21	0.04
BWA	0.89	0.41	0.43	0.85	0.95	0.82	0.5	0.11	0.46
HAI	0.89	0.91	0.0	0.66	0.98	0.95	0.16	0.01	0.22
LIE	0.89	0.89	0.0	0.59	0.93	1.0	0.1	0.01	0.03
LUX	0.89	0.91	0.47	0.73	0.93	0.99	0.84	0.1	0.03
NAM	0.89	0.92	0.25	0.63	0.97	0.99	0.48	0.01	0.2
GR	0.89	0.07	0.0	0.54	0.93	0.97	0.16	0.13	0.02
FR	0.89	0.36	0.35	0.84	0.91	0.93	0.8	0.06	0.06
DR	0.89	0.08	0.08	0.77	0.94	0.95	0.84	0.06	0.09
OV	0.89	0.49	0.15	0.77	0.98	0.98	0.67	0.13	0.1
GE	0.89	0.39	0.36	0.73	0.96	0.99	0.8	0.01	0.11
FL	0.89	0.45	0.07	0.3	0.91	0.65	0.86	0.01	0.08
NH	0.89	0.52	0.19	0.73	0.94	0.97	0.5	0.01	0.0
ZE	0.89	0.46	0.02	0.93	0.9	0.62	0.88	0.05	0.0
UT	0.89	0.69	0.83	0.79	0.95	0.99	0.82	0.08	0.11
ZH	0.89	0.67	0.78	0.56	0.96	0.98	0.8	0.04	0.01
NB	0.89	0.55	0.07	0.76	0.96	0.98	0.7	0.06	0.0
LI	0.89	0.29	0.32	0.78	0.96	0.99	0.04	0.01	0.06

Table 16: R^2, linear regression of endogenous and other factors, Belgium and the Netherlands.

Acknowledgement

I thank the FNRS (Fonds de la recherche scientifique) for the financial support through the EOS project Asterisk (research grant 40007517). I also thank Karim Barigou for his interesting suggestions.

Statements and declarations

The author has seen and agree with the contents of the manuscript and there is no financial interest to report.

References

[1] Cairns A.J.G., Blake D., Dowd K. 2008. Modelling and management of mortality risk: a review. Scandinavian Actuarial Journal, 2008 (2–3), 79-113

[2] Boonen T.J., Li H. 2017. Modeling and forecasting mortality with economic growth: a multipopulation approach. Demography 54(5), 1921–1946.

[3] Brenner M.H. 2005. Commentary: economic growth is the basis of mortality rate decline in the 20th century experience of the United States 1901–2000. International Journal of Epidemiology, 34(6), 1214–1221.

[4] Brouhns N., Denuit M., Vermunt J. 2002. A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics, 31(3), 373-393.

[5] Birchenall J.A. 2007. Economic development and the escape from high mortality. World Development. 35(4), 543–568.

[6] Carmarda C.G. 2019. Smooth constrained mortality forecasting. Demographic research 41, 1091-1130.

[7] Dimai M. 2024. Modeling and forecasting mortality with economic, environmental and lifestyle variables. Decisions in Economics and Finance https://doi.org/10.1007/s10203-024-00434-4

[8] French D., O'Hare C. 2014. Forecasting death rates using exogenous determinants. Journal of Forecasting, 33, 640–650.

[9] Hainaut D. 2018. A neural network analyzer for mortality forecast. ASTIN Bulletin 48(2), 481-508.

[10] Hyndman R. J., Booth H., Yasmeen F., 2013. Coherent mortality forecasting: The product-ratio method with functional time series models. Demography, 50, 261–283.

[11] Li N., Lee R.D., 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. Demography 57, 575-594.

[12] Lee R. D., Carter L. R., 1992. Modeling and forecasting U.S. mortality. Journal of the American Statistical Association, 87, 659–671.

[13] Niu G., Melenberg B., 2014. Trends in mortality decrease and economic growth. Demography, 51(5), 1755–1773.

[14] Oirov T., Terbish G., Dorj N. 2021. B-spline Estimation for Force of Mortality. Mathematics and Statistics, 9(5), 736-743

[15] Pavone F., Legramanti S., Durante D. 2024. Learning and forecasting of age-specific period mortality via B-spline processes with locally-adaptive dynamic coefficients. The Annals of Applied Statistics 18 (3), 1965-1987.

[16] Plat R., 2009. On stochastic mortality modeling. Insurance: Mathematics and Economics, 45(3), 393-404.

[17] Renshaw A.E., Haberman S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics, 38(3), 556–570

[18] Ruhm C.J., 2005. Commentary: Mortality increases during economic upturns. International Journal of Epidemiology, 34, 1206-1211.

[19] Tapia Granados J. A., Roux A.V.D. 2009. Life and death during the Great Depression. Proceedings of the National Academy of Sciences, 106, 17290–17295.

[20] Villegas A.M., Bajekal M., Haberman S. 2025. Modelling mortality by cause of death and socio-economic stratification: an analysis of mortality differentials in England. Decisions in Economics and Finance. https://doi.org/10.1007/s10203-025-00511-2