Longevity Risk Modeling with the Consumer Price Index

Qingxiao Ma* Tim J. Boonen†

July 25, 2023

Abstract

Economic growth has been shown to be an important factor that explains changes in mortality probabilities. Economic growth is commonly measured via the Gross National Product per capita (GDP), but this paper argues that the Consumer Price Index (CPI) is a more natural factor to explain mortality dynamics. It is namely the Consumer Price Index that approximates the affordability of health care, food and housing. We augment the well-known Lee-Carter model with the observable Consumer Price Index factor, and test this model using data from the USA, Canada, Australia, and France. We show that the in-sample model fit of our proposed model improves compared with the Lee-Carter model (either augmented with the Gross National Product factor or not). We also show that the out-of-sample forecasting performance of our proposed model, as measured by the mean squared forecast error, is a considerable improvement. Also, the Lee-Carter model augmented with both the Gross National Product and Consumer Price Index factors performs even better in-sample and out-of-sample.

Keywords: Longevity Risk, Consumer Price Index, Gross Domestic Product, Forecasting.

*China National Health Development Research Center. No. 9 Chegongzhuang Street, Xicheng District, Beijing, China. Email: qingxiao.ma@nhei.cn.

†Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, People's Republic of China. Email: tjboonen@hku.hk.

1 Introduction

Longevity risk modeling is of central importance in life insurance, and the modeling and forecasting of mortality rates are central topics in actuarial science. There is a lack of consensus in the literature on which model is best for mortality rates. A central stochastic mortality risk model is proposed by Lee and Carter (1992), and this is denoted as the by now well-known Lee-Carter model (LC model). An increasing number of researchers have been trying to extend the LC model to improve the accuracy of life expectancy predictions (e.g., Lee and Miller, 2001; Booth et al., 2002; Renshaw and Haberman, 2016; De Jong and Tickle, 2006; Plat, 2009; Lyu et al., 2021).

A key drawback of the LC model is that the future trend in mortality dynamics is captured via a latent trend. This latent trend is obtained via a singular value decomposition (SVD), and its forecasting is generally done via ARIMA models. However, it is precisely this forecasting method that is difficult to justify. Since we do not explicitly know what the latent time trend represents, it is hard to understand its future dynamics.

In the literature, researchers have concentrated on extending the LC model and its related extensions to improve the model fit and forecast performance. Hanewald (2011) finds a strong long-term relationship between the mortality trend and the Gross Domestic Product per capita (GDP) and unemployment rate. Niu and Melenberg (2014) then propose to augment the LC model with the observable factor of economic growth, as measured by the GDP. They show that this improves the model fit, and this is later extended to multiple populations by Boonen and Li (2017) and Cupido et al. (2020) who consider the first principal components of GDP as common factors. Additionally, Seklecka et al. (2019) propose to include an age-dependent factor given by correlation coefficient between GDP and mortality, and this factor is augmented to the mortality model of O'Hare and Li (2012). Using mortality data of sub-populations in Italy, Bozzo et al. (2021) show that it is the GDP level that is a key determinant of mortality dynamics, and not the trend in GDP. They argue that Boonen and Li (2017) may overestimate the effect of the economic crisis on mortality rates because the data set used by Boonen and Li (2017) terminates in 2009 or 2010.

Luo and Xie (2020) show that life expectancy also depends on the level of a country's medical system and the purchasing power of citizens. The purchasing power is an indication of the population's ability to purchase food and other goods for daily consumption. Moreover, we can use GDP to express the development of the economy of a country, but we can alternatively use the consumer price index (CPI) to measure a country's economic stability and people's consumption level. Often, a disaster leads to a substantial increase in inflation (such as currently after the COVID-19 pandemic).

We augment or replace the GDP factor in the model of Niu and Melenberg (2014) by a CPI factor, and we will demonstrate that the model fit is improved both in-sample and out-of-sample by doing so. Using data from USA, Canada, Australia, and France, we first show that CPI is indeed strongly related to the latent time trend in the LC model. Using the new normalization approach introduced by Liu et al. (2019a,b), we provide a mortality model that includes both the log GDP and log CPI, and moreover a latent factor. We show that this model has the best in-sample goodness of fit compared with the Lee-Carter model and the model of Niu and Melenberg (2014). We show the in-sample goodness of fit via the R^2, the adjusted R^2, the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC).

In mortality models with GDP or CPI, the long term dynamics are partially captured by the observable factors. We propose a Vector Error Correction Model (VECM) for the joint modeling of GDP and CPI, and we show that one lag is generally optimal. We demonstrate that the Mean Squared Forecasting Error (MSFE) is smallest for the mortality model with both the log GDP and the log CPI. This paper provides evidence for how observable factors can improve the forecasting of mortality dynamics, which is useful for actuaries and demographers. Moreover, we study the effect of structural breaks in the time-varying factors, which was first proposed by Van Berkum et al. (2016). We find that structural breaks in the time-series do not improve the forecasting performance of our mortality models.

This paper is set out as follows. Section 2 studies the long-term relationship between CPI and mortality. We will introduce the three mortality models related to economic growth in Section 3. In Section 4, we study the performance of the mortality models using data from the USA, Canada, Australia, and France. Section 5 concludes. The appendix provides supplementary findings on the model with structural breaks.

2 Lee-Carter model and CPI

2.1 LC Model

In this section, we first define the well-known Lee-Carter model (Lee and Carter, 1992). The mortality data is available for the years $t = t_0, \ldots, T$ and ages $x = x_0, \ldots, X$, where T denotes the present. In a given sample, the death rate is given by

$$M_{x,t} = \frac{D_{x,t}}{E_{x,t}}, \tag{1}$$

where $D_{x,t}$ is the number of people at age x that are deceased in year t, and $E_{x,t}$ is called the exposure, which is the number of people at age x in year t.

The Lee-Carter (LC) model is defined as follows:

$$\log(M_{x,t}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t}, \tag{2}$$

with $\varepsilon_{x,t}$ being errors with mean zero that are uncorrelated with κ_t. In the classical LC model, the normalizations are given by

$$\sum_{t=t_0}^{T} \kappa_t = 0, \sum_{x=x_0}^{X} \beta_x = 1.$$

However, Liu et al. (2019a,b) propose the following normalization:

$$\sum_{x=x_0}^{X} \alpha_x = 0, \tag{3}$$

$$\sum_{x=x_0}^{X} \beta_x = 1. \tag{4}$$

We decide to use the normalization in (3)-(4), because the restriction $\sum_{t=t_0}^{T} \kappa_t = 0$ may lead to inference pitfalls (Leng and Peng, 2016), and Liu et al. (2019a,b) show that this original normalization may lead to inconsistent estimators.

From (3)-(4), it directly follows that the estimated value of κ_t is equal to $\sum_{x=x_0}^{X} \log(M_{x,t})$. Moreover, we estimate α_x and β_x via Ordinary Least Squares (OLS) as in Liu et al. (2019b). For forecasting, we assume that κ_t has a unit root as shown by Liu et al. (2019b), and follows a random walk with drift:

$$\kappa_t = h + \kappa_{t-1} + \zeta_t, \tag{5}$$

where ζ_t being a white noise error term. Here, the parameter h represents the expected first difference of κ_t.

2.2 CPI and mortality dynamics

Niu and Melenberg (2014) show that GDP is an important macro-economic factor that can help to model and forecast log death rates. In this paper, we argue that CPI is a better indicator to measure a country's economic situation and the living standards of its citizens. We are interested in the effect of CPI on log death rates and whether CPI will be a better factor than GDP to explain the trend of future mortality dynamics. Furthermore, we will use both CPI and GDP to estimate and forecast mortality. We define the two economic growth factors as follows:

$$c_t = \log(CPI_t) - \frac{1}{T - t_0 + 1} \sum_{t'=t_0}^{T} \log(CPI_{t'}), \text{ and} \tag{6}$$

$$G_t = \log(GDP_t) - \frac{1}{T - t_0 + 1} \sum_{t'=t_0}^{T} \log(GDP_{t'}), \tag{7}$$

where the log refers to the natural logarithm. Here, the factors c_t and G_t represent the zero-normalized, time-t values of the log CPI and log GDP, respectively.

We use mortality data for females[1] in the countries USA, Canada, Australia, and France, and this data set is available in the Human Mortality Database (HMD). We have data for ages $x = 0, \ldots, 110$ and time periods $t = 1970, \ldots, 2018$, so that we denote $x_0 = 0$, $X = 110$, $t_0 = 1970$, and $T = 2018$. The data on the CPI and GDP are obtained from the OECD database[2,3]. The GDP data is corrected for inflation to the 2021 price level and under the 2021 purchasing power parity (PPP).

[1] In this paper, we only study the female populations, but we get qualitative similar results if we were to use the data for males or both sexes combined.

[2] OECD (2021), Inflation (CPI) (indicator). doi:10.1787/eee82e6e-en (accessed on June 4, 2022)

[3] OECD (2021), Gross domestic product (GDP) (indicator). doi: 10.1787/dc2f7aec-en (Accessed on June 4, 2022)

Variable	Countries	p-value ($r \leq 0$)	p-value ($r \leq 1$)
c_t	USA	0.0449**	0.0456**
	Canada	0.0041***	0.5046
	Australia	0.0090***	0.7594
	France	0.0133**	0.1060
G_t	USA	0.0010***	0.0118**
	Canada	0.0010***	0.0010***
	Australia	0.0022***	0.0822*
	France	0.0132**	0.0438**

*** indicates p-value smaller than 0.01 ** indicates p-value smaller than 0.05, and * indicates p-value smaller than 0.1.

Table 1: Johansen test for both c_t and G_t in relation with κ_t from the LC model for all selected countries.

Countries	p-value ($r \leq 0$)	p-value ($r \leq 1$)
USA	0.0010***	0.0011***
Canada	0.0027***	0.0929*
Australia	0.0011***	0.0756*
France	0.0198**	0.0747*

*** indicates p-value smaller than 0.01 ** indicates p-value smaller than 0.05, and * indicates p-value smaller than 0.1.

Table 2: Johansen test for the time series G_t and c_t for all selected countries.

Next, we test the long-run relationship between log CPI and mortality. To do so, we use the Johansen cointegration test on κ_t in the LC model and the zero-normalized value of the log CPI (c_t), where the method to calculate κ_t is given by Equation (2) and the p-values of the Johansen test are given in Table 1.

In Table 1, the null hypotheses of $r \leq 0$ and $r \leq 1$ state that there is no cointegration and one cointegration vector exists, respectively. We find that all outcomes support rejection of the null hypothesis of $r \leq 0$ under a 5% confidence level. This means that we reject the hypothesis that there does not exist any cointegration. Thus, we find that there exists a long-run relationship between c_t (from Equation (6)) and the log death rates. This inspires our central question whether CPI can explain the mortality trend better than GDP and whether the model with both factors can explain the mortality trend better than a model with one factor.

Table 1 also confirms the findings of Niu and Melenberg (2014) that the log GDP factor G_t is cointegrated with the mortality trend κ_t, under a 1% confidence level for USA, Canada and Australia, and under a 5% confidence level for France. Thus, both the log GDP and log CPI are good candidates to study their impact on mortality. Table 2 provides evidence that the factor c_t and G_t are also cointegrated, meaning that both factors may jointly pick up a common economic growth trend.

3 Three mortality models

3.1 LC-GDP model

Niu and Melenberg (2014) use the GDP as a macro-economic risk factor to estimate and forecast the log death rate; the LC-GDP model is the first stochastic model of mortality with economic growth. The LC-GDP model is given by

$$\log(M_{x,t}) = \alpha_x + \beta_x \kappa_t + \gamma_x G_t + \varepsilon_{x,t}, \tag{8}$$

with $\varepsilon_{x,t}$ being errors with mean zero that are uncorrelated with κ_t and G_t, where $M_{x,t}$ is defined in (1) and G_t is defined in (7). We impose the normalization constraints (3)-(4), and moreover $cov(\kappa_t, G_t) = 0$ and $(\kappa_{t_0}, \ldots, \kappa_T) \neq \mathbf{0}$. Here, the main contrast with the LC model is that the LC-GDP model considers G_t as a risk factor that drives the long-term trend in mortality dynamics instead of the latent factor κ_t.

The normalization constraints $cov(\kappa_t, G_t) = 0$ and $(\kappa_{t_0}, \ldots, \kappa_T) \neq \mathbf{0}$ are based on Theorem 1 in Niu and Melenberg (2014). Without these constraints, it follows that the model parameters are not identified. The proof of this identification statement in Niu and Melenberg (2014) follows from arguments in Kuang et al. (2008a), Kuang et al. (2008b), and Nielsen and Nielsen (2014). Moreover, Boonen and Li (2017) also provide a similar identification result.

Our estimation method is as follows. First, we estimate the parameter γ_x via OLS, where γ_x is the coefficient of G_t. Second, we estimate κ_t as $\sum_{x=x_0}^{X}(log(M_{x,t}) - \gamma_x G_t)$, and then the parameters α_x and β_x are estimated via OLS as in the LC model, but we now use $\log(M_{x,t}) - \gamma_x G_t$ as the dependent variable.[4] Note that since we use the normalization of Liu et al. (2019a,b) in the LC-GDP model, the estimation in this second step is different compared with Boonen and Li (2017).

3.2 LC-CPI model

In this subsection, we introduce our first new model of this paper: the LC-CPI model. This model considers CPI as an observable risk factor that captures the long-term trend of mortality, instead of GDP. The LC-CPI model is given as follows:

$$\log(M_{x,t}) = \alpha_x + \beta_x \kappa_t + \tau_x c_t + \varepsilon_{x,t}, \tag{9}$$

with $\varepsilon_{x,t}$ being errors with mean zero that are uncorrelated with κ_t and c_t, and with normalization constraints (3)-(4), $cov(\kappa_t, c_t) = 0$, and $(\kappa_{t_0}, \ldots, \kappa_T) \neq \mathbf{0}$. Here, $M_{x,t}$ is defined in (1). Compared with the LC model, we add the factor c_t which is defined in (6).

Both the LC-GDP and the LC-CPI model are estimated in two steps. In the LC-CPI model, the value of τ_x is interpreted as the sensitivity of $\log(M_{x,t})$ at age x to variations in c_t. If we were to reverse the estimation steps, then the constraint $cov(\kappa_t, c_t) = 0$ is typically violated, and also we then implicitly

[4]Orthogonality of κ_t with G_t is imposed to enable the two-step estimation. Moreover, if κ_t is not orthogonal to G_t, the first-step estimates of γ_x have no immediate interpretation as the sensitivity of the log death rate at age x to variations in G_t. Similar arguments are used for the LC-CPI and LC-GDP-CPI models, which are defined later in Sections 3.2 and 3.3.

impose the constraint $\sum_{x=x_0}^{X} \tau_x = 0$. It is an undesirable constraint because it implies that there exists some age groups that have an opposite sensitivity to c_t than the other age groups.

3.3 LC-GDP-CPI model

In this subsection, we merge the LC-GDP and LC-CPI models and denote this model as the LC-GDP-CPI model. It is defined as follows:

$$\log(M_{x,t}) = \alpha_x + \beta_x \kappa_t + \gamma_x G_t + \tau_x c_t + \varepsilon_{x,t}, \tag{10}$$

with $\varepsilon_{x,t}$ being errors with mean zero that are uncorrelated with κ_t, G_t and c_t. In this model, we consider both the variations of GDP and CPI that simultaneously affect mortality rates. The LC-GDP-CPI model is the combination of the LC-GDP model (Section 3.2) and the LC-CPI model (Section 3.3). We impose the same normalization constraints as in the LC-GDP and LC-CPI models, which are (3)-(4), $(\kappa_{t_0}, \ldots, \kappa_T) \neq \mathbf{0}$, $cov(\kappa_t, c_t) = 0$ and $cov(\kappa_t, G_t) = 0$. The estimation technique is similar to the LC-GDP or LC-CPI model. We first apply an OLS estimation of the parameters γ_x and τ_x. Thereafter, we estimate κ_t as $\sum_{x=x_0}^{X}(log(M_{x,t}) - \gamma_x G_t - \tau_x c_t)$, and the parameters α_x and β_x are estimated via an OLS as in the LC model, but with $\log(M_{x,t}) - \gamma_x G_t - \tau_x c_t$ as the dependent variable.

3.4 Modeling of time series

In this section, we specify the time-series models used for forecasting. We first focus on the model for the macro-economic factors (c_t, G_t), and then we focus on the model for κ_t, which is assumed to be orthogonal to the macro-economic factors that are included in the particular mortality model.

We use the Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test to test for stationarity of c_t and G_t. The null hy-

Variable	Countries	ADF test	KPSS test
c_t	USA	-3.2656^{***}	1.0834^{***}
	Canada	-3.8451^{***}	1.1135^{***}
	Australia	-3.8767^{***}	1.1186^{***}
	France	-4.7172^{***}	1.0981^{***}
G_t	USA	-2.8735^{***}	1.1212^{***}
	Canada	-3.1375^{***}	0.9945^{***}
	Australia	-2.2861^{**}	0.9883^{***}
	France	-3.0475^{***}	1.0537^{***}

Notes: The critical values for the null hypothesis in the ADF test are -1.6115 for $p < 0.1$, -1.9468 for $p < 0.05$, and -2.6136 for $p < 0.01$. The critical values for the null hypothesis in the KPSS test are 0.1190 for $p < 0.1$, 0.1460 for $p < 0.05$, and 0.2160 for $p < 0.01$.

Table 3: Test statistics of the two unit root tests for c_t and G_t.

pothesis in Table 3 is that the variable has no unit root in KPSS test and it does have a unit root in ADF test. In other words, under the null hypotheses, the variable is stationary in the KPSS test and it is not stationary in the ADF test. Table 3 shows that both c_t and G_t reject the null hypotheses under a 5% confidence level in both the ADF and KPSS tests; this is a contradictory conclusion. In such case, we chose to rely on the conclusion of the KPSS test, because Kwiatkowski et al. (1992) argue that there exists a unit root in most economic time series. Thus, G_t and c_t both have a unit root.

For the LC-GDP and LC-CPI models, we prefer to use the same model formulations as in Niu and Melenberg (2014) and Boonen and Li (2017), respectively. We model and forecast the log GDP factor G_t and the log CPI factor c_t with a unit root as random walks with drift, which coincides with the time-series model in (5).

Next, we aim to select an appropriate time-series model for the log CPI and log GDP factors (c_t, G_t) in the LC-GDP-CPI model. Since c_t and G_t are both shown to have a unit root, we propose modeling $y_t := (c_t, G_t)$ with a VECM model (see, e.g., Agbonlahor, 2014). The VECM$(p-1)$ model for $\Delta y_t := y_t - y_{t-1}$ is defined as follows:

$$\Delta y_t = c + dt + AB' y_{t-1} + \sum_{i=1}^{p-1} \Phi_i \Delta y_{t-i} + \nu_t, \tag{11}$$

where d is the constant time-trend parameter, and A and B are parameters that represent the adjustment speed and the cointegration matrix, respectively. Moreover, ν_t is a Gaussian i.i.d. error term with zero mean.

By constructing a VECM model, we can account for the long-run relationship between the log GDP and log CPI factors. In order to select the number of lags in the VECM model of y_t, we first use the VAR(p) model to find the optimal number of lags p, and then we construct the VECM$(p-1)$ model. The VAR(p) model is given by:

$$c_t = \lambda_c + \sum_{i=1}^{p} \phi_{c,c,i} c_{t-i} + \sum_{i=1}^{p} \phi_{c,G,i} G_{t-i} + \zeta_{t,1}, \tag{12}$$

$$G_t = \lambda_G + \sum_{i=1}^{p} \phi_{G,c,i} c_{t-i} + \sum_{i=1}^{p} \phi_{G,G,i} G_{t-i} + \zeta_{t,2}, \tag{13}$$

where $\zeta_{t,\cdot} \overset{i.i.d.}{\sim} N(0, \Sigma)$. Here, $\phi_{a,b,i}$ represents the coefficient for the lagged dependent variable $a \in c, G$ and the lagged independent variable $b \in c, G$ at lag i. We need to select the number of lags, i.e., select the value of p in (12)-(13). To do so, we use the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which are given by:

$$AIC = -2\log(L) + 2k, \tag{14}$$

$$BIC = -2\log(L) + \log(n) \cdot k, \tag{15}$$

where L is the maximum value of the likelihood function for the model assuming Gaussianity and homoscedasticity over time of all error terms, k is the number of estimated parameters in the model, and n is the size of the sample. The model that is endowed with the lowest AIC or BIC is then subsequently selected as the best model. We refer to Yang (2005) for an overview of the differences between the AIC and BIC values.

Countries	Number of lags p	AIC	BIC
USA	1	-382.44	-372.45
	2	-409.31*	-392.94*
	3	-398.73	-376.17
	4	-384.29	-355.79
	5	-376.82	-342.61
Canada	1	-366.21	-356.23
	2	-408.03*	-391.65*
	3	-393.47	-370.92
	4	-376.07	-347.57
	5	-358.91	-324.69
Australia	1	-349.30	-339.31
	2	-375.26	-358.88*
	3	-380.38*	-357.83
	4	-368.42	-339.92
	5	-365.24	-331.02
France	1	-367.26	-357.28
	2	-419.14*	-402.76*
	3	-405.49	-382.94
	4	-393.12	-364.62
	5	-386.42	-352.20

* indicates the smallest value of AIC or BIC in that country.

Table 4: Model selection criteria AIC and BIC for the VAR(p) model of (c_t, G_t), where p is the number of lags.

As shown in Table 4, the AIC value is the smallest when there are two lags ($p = 2$), except for Australia that has the smallest AIC value when $p = 3$. Moreover, the BIC value is smallest when $p = 2$ in all selected countries. Therefore, we select $p = 2$ as the optimal number of lags in the VAR(p) model, and thus we will use the VECM(1) model to estimate and forecast (c_t, G_t).

To select the distribution of κ_t in the LC-GDP, LC-CPI, and LC-GDP-CPI models, we first perform the ADF and KPSS tests. The test statistics are shown in Tables 5 and 6. We find that κ_t is non-stationary based on both the ADF test and the KPSS test, and $\Delta\kappa_t$ is stationary in most cases. We model κ_t as an ARIMA(p,1,q) model, given by:

$$\Delta\kappa_t = \lambda_{\Delta\kappa} + \sum_{i=1}^{p} \omega_i \Delta\kappa_{t-i} + \sum_{j=1}^{q} \psi_j \epsilon_{t-j} + \zeta_t, \tag{16}$$

where ζ_t is a white noise error term. We select p and q based on the AIC and BIC model selection criteria.

Test	Models	USA	Canada	Australia	France
	LC	4.3004	6.8392	4.5817	6.2588
ADF	LC-GDP	0.0818	-0.7533	0.3277	-0.6591
test	LC-CPI	0.1740	-0.4282	0.4086	0.0624
	LC-GDP-CPI	0.1403	-0.7377	0.3509	-0.7435
	LC	0.5565***	0.7852***	0.3567***	0.1406*
KPSS	LC-GDP	0.3955***	0.5488***	0.6770***	0.9842***
test	LC-CPI	0.6429***	1.0632***	1.0655***	1.0756***
	LC-GDP-CPI	0.5618***	0.1710***	0.1643***	0.1481**

Notes: The critical values for the null hypothesis in the ADF test are -1.6115 for $p < 0.1$, -1.9468 for $p < 0.05$, and -2.6136 for $p < 0.01$. The critical values for the null hypothesis in the KPSS test are 0.1190 for $p < 0.1$, 0.1460 for $p < 0.05$, and 0.2160 for $p < 0.01$.
*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Table 5: Test statistics for the unit root test for κ_t.

Test	Models	USA	Canada	Australia	France
	LC	-4.6912***	-3.4294***	-6.7462***	-5.1377***
ADF	LC-GDP	-6.5518***	-5.7238***	-9.9633***	-7.1384***
test	LC-CPI	-7.5109***	-6.2875***	-9.0265***	-5.7209***
	LC-GDP-CPI	-7.2538***	-5.4602***	-9.6840***	-6.9622***
	LC	0.0813	0.0733	0.0295	0.0460
KPSS	LC-GDP	0.0653	0.1014	0.0202	0.1375*
test	LC-CPI	0.0613	0.1668**	0.0283	0.1971**
	LC-GDP-CPI	0.0629	0.0884	0.0255	0.0905

Notes: The critical values for the null hypothesis in the ADF test are -1.6114 for $p < 0.1$, -1.9469 for $p < 0.05$, and -2.6148 for $p < 0.01$. The critical values for the null hypothesis in the KPSS test are 0.1190 for $p < 0.1$, 0.1460 for $p < 0.05$, and 0.2160 for $p < 0.01$.
*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Table 6: Test statistics for the unit root test for $\Delta\kappa_t$.

The model selection criteria *AIC* and *BIC* are shown in Table 7, and in case of a conflicting model selection we chose to follow the recommendation according to the *BIC*. We find that κ_t is best modelled as a random with a drift in the USA in all mortality models. Moreover, for all mortality models, we choose the ARIMA$(2, 1, 2)$ model in Canada. In Australia, we select the ARIMA$(0, 1, 1)$ model for the LC-GDP and LC-GDP-CPI models and ARIMA$(2, 1, 2)$ for the LC-CPI model. Finally, in France, we select the ARIMA$(1, 1, 2)$ model in the LC-GDP model and the ARIMA$(1, 1, 1)$ model in the LC-CPI and LC-GDP-CPI models.

Model	Countries	AIC	BIC

LC-GDP	USA	(0,0)	(0,0)
	Canada	(2,2)	(2,2)
	Australia	(3,2)	(0,1)
	France	(1,2)	(1,2)
LC-CPI	USA	(2,2)	(0,0)
	Canada	(3,2)	(2,2)
	Australia	(2,2)	(2,2)
	France	(1,1)	(1,1)
LC-GDP-CPI	USA	(0,0)	(0,0)
	Canada	(2,2)	(2,2)
	Australia	(3,2)	(0,1)
	France	(1,1)	(1,1)

Table 7: The optimal number of lags of κ_t in the ARIMA(p,1,q) model.

4 Estimation and performance

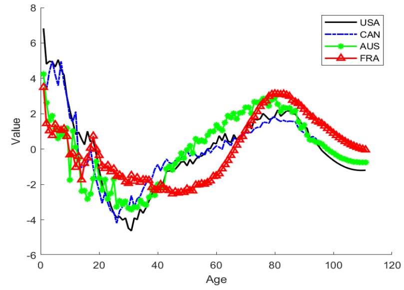
In this section, we discuss the estimation and show the in-sample and out-of-sample performance of all four mortality models that we study in this paper.

4.1 Estimation

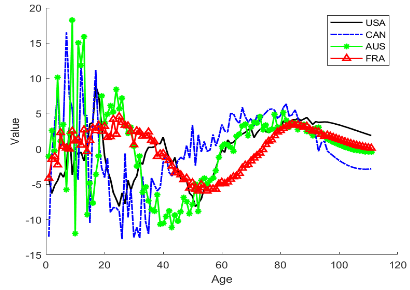
With the normalization of Liu et al. (2019a,b), α_x represents the deviations from the average log death rate of $\log(M_{x,t})$ for each age x.

We expect that the estimated parameters α_x in the models including the economic growth factors vary more than estimates of α_x in the LC model, because κ_t in the LC model is latent and selected as the optimal age-invariant factor while the log CPI and the log GDP are observable age-invariant factors. Indeed, the estimates of α_x, displayed in Figure 1, are in line with our intuition.

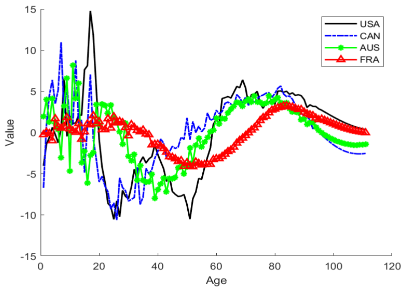
Furthermore, since the CPI is highly correlated to the GDP, we consider that the estimates of α_x in the LC-GDP model are similar as the ones in the LC-CPI model, and Figures 1(b) and 1(c) verify this. Also, we expect that the estimates of α_x for the populations in the LC model, displayed in Figure 1(a), should be similar to those in the other models. For French children aged between 0 and 17, this figure shows substantial differences compared to other countries for the LC-GDP and LC-CPI models, whereas the difference in the LC-GDP-CPI model is smaller (see Figure 1(d)).



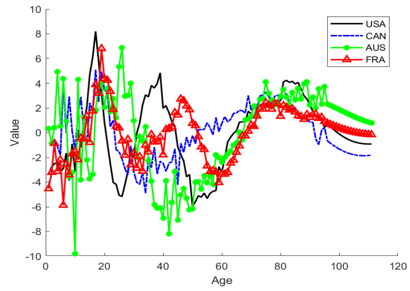
(a) LC model.



(b) LC-GDP model.

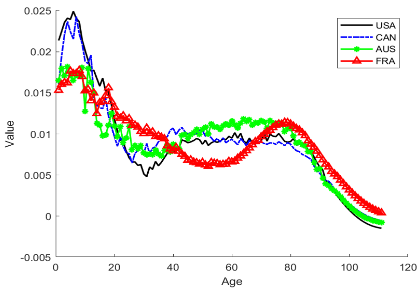


(c) LC-CPI model.

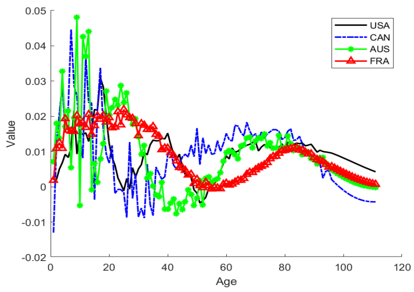


(d) LC-GDP-CPI model.

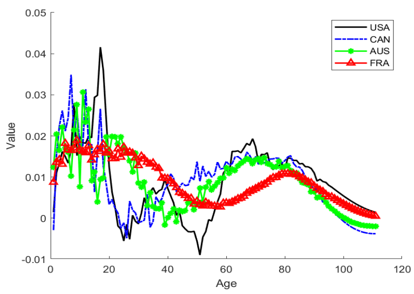
Figure 1: The estimates of α_x in all selected countries based on the four mortality models.



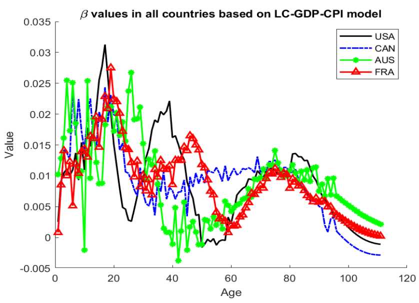
(a) The estimates of β_x in all selected countries based on the LC model.



(b) The estimates of β_x in all selected countries based on the LC-GDP model.

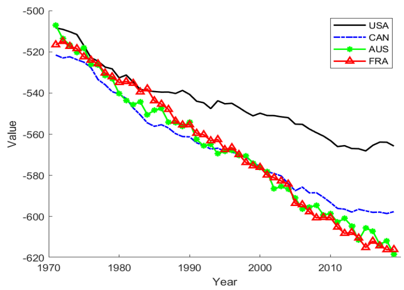


(c) The estimates of β_x in all selected countries based on the LC-CPI model.

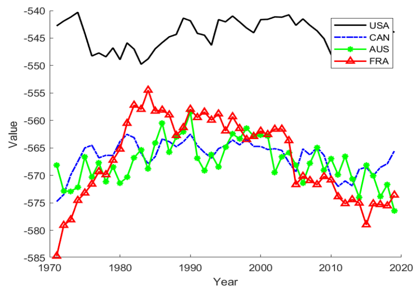


(d) The estimates of β_x in all selected countries based on the LC-GDP-CPI model.

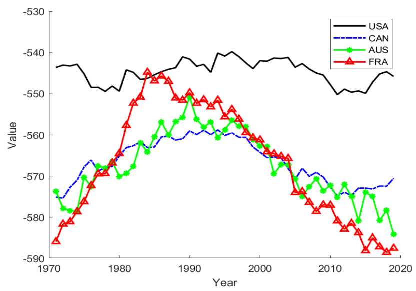
Figure 2: The estimates of β_x in all selected countries based on the four mortality models.



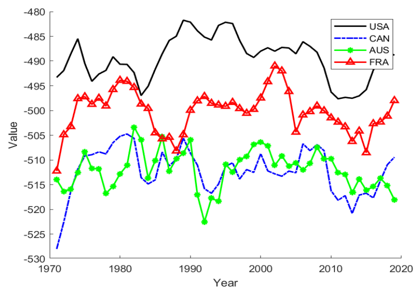
(a) The estimates of κ_t in all selected countries based on the LC model.



(b) The estimates of κ_t in all selected countries based on the LC-GDP model.



(c) The estimates of κ_t in all selected countries based on the LC-CPI model.

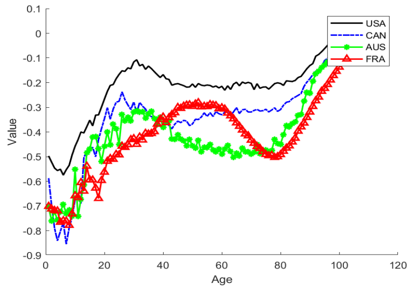


(d) The estimates of κ_t in all selected countries based on the LC-GDP-CPI model.

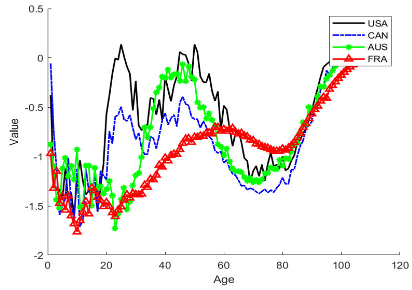
Figure 3: The estimates of κ_t in all selected countries based on the four mortality models.

The parameter β_x represents the age-dependent sensitivity of $\log(M_{x,t})$ to the latent factor κ_t, and we display the estimates of β_x in Figure 2. For the LC model, the parameter β_x captures the sensitivity to a long-term time trend of mortality dynamics. For the other mortality models, we do not find a clear common pattern in the estimates of β_x. Due to our normalization (3)-(4) as in Liu et al. (2019a,b), the estimate of κ_t is equal to the aggregate log death rate in year t. Generally, the aggregate log death rates have a decreasing trend, and this decreasing trend may at least partially be due to economic growth. The estimates of κ_t in the LC-GDP model (Figure 3(b)), the LC-CPI model (Figure 3(c)) and the LC-GDP-CPI model (Figure 3(d)) do no longer display a strong trend as in the LC model (Figure 3(a)). So, although κ_t is non-stationary (recall Table 5), κ_t displays a weaker trend than the macro-economic factors that are considered in the model. We interpret this as that the economic growth factors are suitable to predict the decreasing trend in the mortality rates.

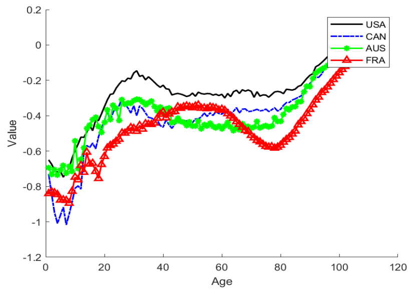
The parameters γ_x and τ_x represent the age-dependent impact to $\log(M_{x,t})$ of G_t and c_t, respectively. Generally, when we consider a model including one economic growth factor (the LC-GDP and LC-CPI model), the impact of the economic growth factor on log death rates is negative and decreases as age increases (see Figures 4(a) and 4(c)). In these two models, this indicates that the economic growth factor always negatively influences mortality. For the LC-GDP-CPI model, this impact is less clear. We see that c_t often has a positive



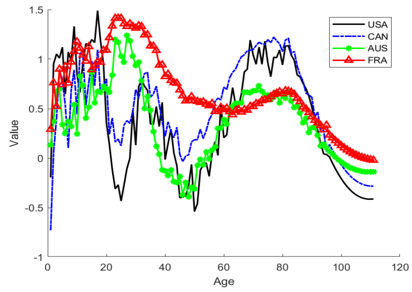
(a) The estimates of γ_x for all selected countries based on the LC-GDP model.



(b) The estimates of γ_x for all selected countries based on the LC-GDP-CPI model.



(c) The estimates of τ_x for all selected countries based on the LC-CPI model.



(d) The estimates of τ_x for all selected countries based on the LC-GDP-CPI model.

Figure 4: The estimates of γ_x and τ_x for all selected countries based on all four mortality models.

coefficient τ_x. The net impact of CPI on log death rates can however still be negative because GDP and CPI are positively correlated and the estimates of γ_x are generally negative.

4.2 In-sample performance

We study the in-sample goodness of fit of our models. To do so, we first compute the R^2 for every model. For a given model, this is given by

$$R^2 = 1 - \frac{\sum_{x=x_0}^{X} \sum_{t=t_0}^{T} (\log(M_{x,t}) - \widehat{\log(M_{x,t})})^2}{\sum_{x=x_0}^{X} \sum_{t=t_0}^{T} (\log(M_{x,t}) - \frac{1}{T-t_0+1} \sum_{t'=t_0}^{T} \log(M_{x,t'}))^2}, \tag{17}$$

where $\log(M_{x,t})$ is the observed value of the log death rate at age x and year t, and $\widehat{\log(M_{x,t})}$ gives the model-based estimated value of the log death rate at age x and year t. The higher the value of R^2, the better the goodness of fit.

Table 8 displays the values of R^2 for all four different mortality models and the four selected countries. The LC-GDP-CPI model yields the largest value of R^2, and the LC-CPI model yields the second largest value in all countries, except for the USA. Thus, the mortality models including CPI have a larger goodness-of-fit than the LC-GDP model, providing some empirical evidence that CPI is a better macro-economic factor to explain mortality dynamics than

Model	USA	Canada	Australia	France
LC	0.8933	0.8691	0.8552	0.9397
LC-GDP	0.9067	0.8757	0.8671	0.9509
LC-CPI	0.9080	0.8802	0.8689	0.9516
LC-GDP-CPI	0.9344*	0.8832*	0.8745*	0.9531*

* indicates the maximal value of R^2 for that country among the four mortality models.

Table 8: The R^2, defined in (17), for all four mortality models and all four selected countries.

GDP. Furthermore, Table 9 shows proportion of variance explained by the time-

Country	Parameters	LC model	LC-GDP model	LC-CPI model
USA	κ_t	0.8933	0.0284	0.0368
	G_t		0.8782	
	c_t			0.8712
	Total	0.8933	0.9067	0.9080
Canada	κ_t	0.8691	0.0199	0.0561
	G_t		0.8558	
	c_t			0.8241
	Total	0.8691	0.8757	0.8819
Australia	κ_t	0.8552	0.0263	0.0789
	G_t		0.8408	
	c_t			0.7900
	Total	0.8552	0.8671	0.8689
France	κ_t	0.9397	0.0648	0.1887
	G_t		0.8861	
	c_t			0.7630
	Total	0.9397	0.9509	0.9516

Table 9: The proportion of variance explained by the time-varying factors in the LC model, the LC-GDP model, and the LC-CPI model.

varying factors κ_t, G_t and c_t. Both factors G_t and c_t substantially reduce the goodness of fit of the latent factor κ_t. This is consistent with our above consideration that economic growth factors can explain $\log(M_{x,t})$, and it verifies that the economic growth factor is indeed correlated with the mortality trend.

It is well-known that the value of R^2 will go up when the models are nested. The mortality models here are nested, as the LC model is a special case of all other models and the LC-GDP and LC-CPI models are special cases of the LC-GDP-CPI model. Therefore, it is expected that the LC-GDP-CPI model has the largest R^2 value compared with other models. Therefore, we also consider the adjusted R^2, denoted by R^2_{adj}, to determine the in-sample model performance. The R^2_{adj} is given by

$$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-k}, \tag{18}$$

where R^2 is defined in (17), n is the sample size, the value of k is the number of independent variables in the model. Here, R^2_{adj} includes a penalty term for the number of independent variables.

Model	USA	Canada	Australia	France
LC	0.8930	0.8688	0.8548	0.9396
LC-GDP	0.9063	0.8752	0.8666	0.9508
LC-CPI	0.9077	0.8799	0.8685	0.9515
LC-GDP-CPI	0.9340*	0.8826*	0.8739*	0.9530*

* indicates the maximal value of R^2_{adj} in that country.

Table 10: The R^2_{adj} based on the four mortality models in the four selected countries.

Table 10 displays the values of R^2_{adj} for all selected countries, and we find that R^2_{adj} yields the same conclusion as the values of R^2 in Table 8. The LC-GDP-CPI model still has the best goodness of fit.

Besides R^2_{adj}, we also study the AIC and BIC values for model selection; see (14)-(15) for the definition. Table 11 shows the AIC and BIC values for

Criteria	Model	USA	Canada	Australia	France
AIC	LC	-15890	-13092	-11685	-15842
	LC-GDP	-16541	-13316	-12012	-17353
	LC-CPI	-16665	-14025*	-12292	-17419
	LC-GDP-CPI	-18274*	-14023	-12951*	-17428*
BIC	LC	-15286	-12488	-11080	-15238
	LC-GDP	-15633	-12408	-11104	-16445
	LC-CPI	-15757	-13117*	-11384	-16511*
	LC-GDP-CPI	-17063*	-12812	-11739*	-16217

* indicates the minimal value of AIC or BIC in that country.

Table 11: The AIC and BIC values for the four mortality models in the four selected countries.

all four mortality model and all selected countries. From this table, it follows that the in-sample goodness-of-fit is the best for the LC-GDP-CPI model. The LC-CPI model has the smallest AIC and BIC in some cases, but the differences between the LC-CPI and LC-GDP-CPI models are not large.

We next use the mean squared forecast error (MSFE) to study the out-of-sample forecasting performance. We fit the mortality model on the data with year $t \in \{t_0, \ldots, u\}$ for jump-off year $u \in \{t_0, \ldots, T-1\}$. Then, we forecast the mortality rates for the period $u+1, \ldots, T$, and denote by $\widehat{\log(M_{x,t})}$ the forecasted value of the log death rate. The MSFE is then given by:

$$\text{MSFE}(u) = \frac{1}{(T-u)(X-x_0+1)} \sum_{t=u+1}^{T} \sum_{x=x_0}^{X} (\log(M_{x,t}) - \widehat{\log(M}_{x,t}))^2, \quad (19)$$

with $u \in \{t_0, \ldots, T-1\}$ the jump-off year, and where $\log(M_{x,t})$ is the observed log death rate. In Table 12, u is equal to the year 2009, so that the forecasting is evaluated based on the last 9 years in the sample.

Model	USA	Canada	Australia	France
LC	0.0175	0.0314*	0.0375	0.0148
LC-GDP	0.0181	0.0342	0.0351	0.0211
LC-CPI	0.0158	0.0345	0.0338*	0.0185
LC-GDP-CPI	0.0125*	0.0332	0.0345	0.0127*

* indicates the minimal value of MSFE in that country compared with other models.

Table 12: The MSFE for all selected countries based on the four different mortality models with $u = 2009$, where MSFE is defined in Table 19.

The MSFE values in Table 12 indicate that the LC-GDP-CPI model performs best in out-of-sample forecasting, and the LC-CPI model performs better than the LC-GDP model in all selected countries, except in Canada. So, including the CPI as one economic growth factor instead of the GDP reduces the MSFE, and it can therefore improve the mortality forecasts.

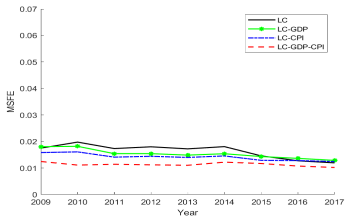
Table 12 only considers the jump-off year $u = 2009$. We next consider multiple jump-off years, because we wish to obtain more evidence to check whether the new models (LC-CPI and LC-GDP-CPI models) persistently perform better than the LC and LC-GDP models. Therefore, we also fit the four mortality models for all selected countries up to the jump-off years $u \in \{2009, \dots, 2017\}$. Then, we used data range from 1970 to year u to find the estimates of the parameters, and then we used these estimates to forecast $\log(M_{x,t})$ for time t between $u+1$ and 2018, and compute the MSFE. In this way, we assess whether the forecasting accuracy is stable in the mortality models.

Figure 5 shows the value of MSFE based on different models and different jump-off year in all selected countries. The LC-GDP-CPI model gives the minimal MSFE in all selected countries, Moreover, the model with one economic growth factor cannot forecast mortality better than the LC model in Canada or France. For France this poor performance may be due to the weak cointegration relationship between c_t and κ_t (recall Table 2), and it implies that one economic growth factor is not enough to explain the mortality trend. On the other hand, in most cases, the model including both the GDP and the CPI can effectively reduce the forecasting error.

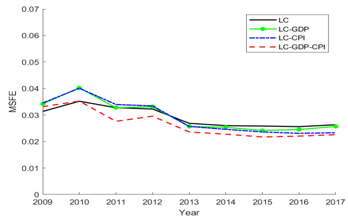
4.3 Structural breaks

In Section 3, we argue that a main trend in mortality is given by the macroeconomic factors, or κ_t in the LC model. We forecast these time-varying factors via an ARIMA(p,1,q) or VECM model. This implies that the mortality dynamics are non-stationary and there may be a non-linear mortality trend. However, one may question whether structural breaks may occur in this trend. With a structural break, the trend in the time-varying factors is assumed to be an $ARIMA(p, 1, q)$ model on subintervals of the time horizon. To test for structural breaks, we adopt the method of Van Berkum et al. (2016). This method was inspired by the ideas of Zeileis et al. (2003) and Li et al. (2015).

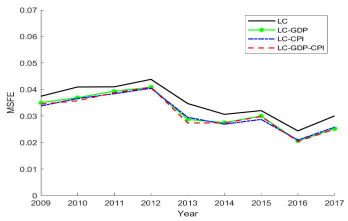
For selecting the number of structural breaks and the break points themselves, Van Berkum et al. (2016) propose two methods. The first method is to use an F-test as proposed by Bai and Perron (1998, 2003). The second method



(a) The comparison of MSFE in USA based on the four mortality models.



(b) The comparison of MSFE in Canada based on the four mortality models.



(c) The comparison of MSFE in Australia based on the four mortality models.

(d) The comparison of MSFE in France based on the four mortality models.

Figure 5: The comparison of MSFE in the four selected countries and based on all four different mortality models.

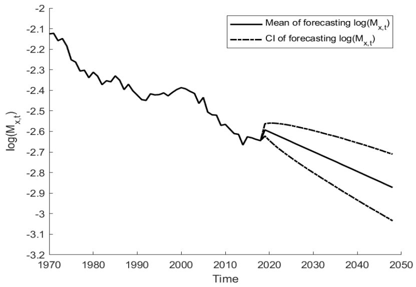
is to select the one with the lowest BIC value. We use the second method because the estimator based on the optimal BIC value is consistent with the true number of break points (see Yao, 1988). Hence, we use the BIC value as our main criterion, and the corresponding outcomes are shown in Table 13 in Appendix A.

We next use the optimal break points, and determine the out-of-sample forecast performance via the MSFE. It is clear that the MSFE values, given in Figure 8 in Appendix A, in the models with break points are larger in most cases than those in the models without structural breaks. Thus, since we want to select the model with the best out-of-sample performance, we will no longer consider the mortality models with structural breaks, and continue to focus on the full-period sample for forecasting mortality.

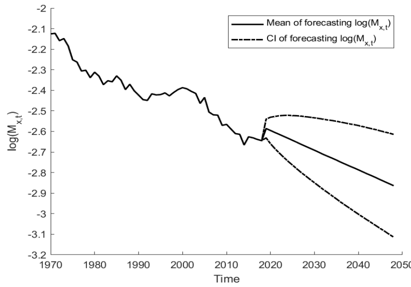
5 Forecasting mortality

In this section, we provide the forecasts of our four mortality models in the USA. We generate 10,000 Monte Carlo simulations, and forecast $\log(M_{x,t})$ at age $x = 85$ and the period life expectancy (LE) at birth. The uncertainties follow from the random projections of the time-dependent factors. The period LE is defined by, e.g., Shkolnikov et al. (2011). Recall that the sample is from 1970 to 2018, and we forecast the two variables for 30 years with the forecasting window from 2019 to 2048.

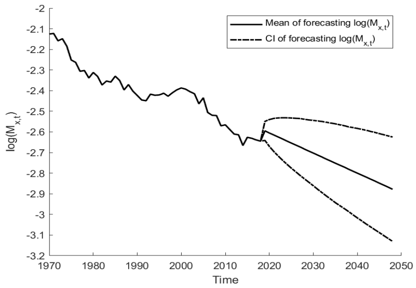
Figure 6 shows the forecasts of $\log(M_{x,t})$ for age $x = 85$ in USA under the



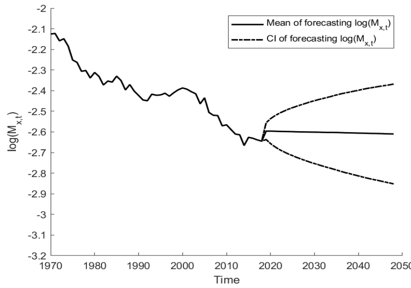
(a) LC



(b) LC-GDP



(c) LC-CPI

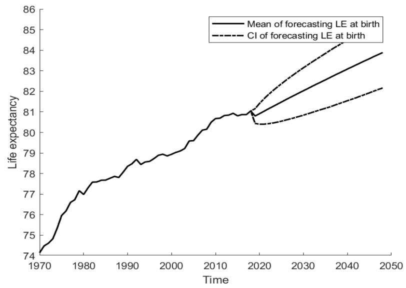


(d) LC-GDP-CPI

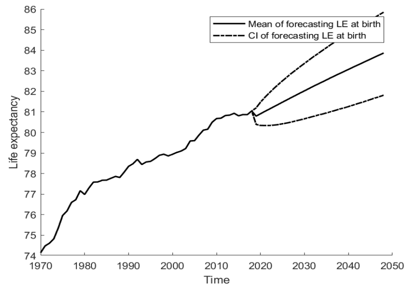
Figure 6: Forecasting $\log(M_{x,t})$ for $x = 85$ in the USA based on the four mortality models.

four mortality models. For instance, the LC model yields an decreasing trend in the forecasts of the log death rate, and the 95%-confidence interval (CI) reaches the interval $[-2.71, -3.03]$ at year 2048. All four mortality models, including the ones with only one macro-economic factor, give an optimistic projection of future mortality rates as they display a decreasing trend, but also show a small jump upwards in the first period of the forecasting window. The LC-GDP-CPI model is the most conservative model with the widest confidence intervals, and yields only a small expected decrease in the log death rates of an 85-year old in the forecasting window. Djeundje et al. (2022) show that many demographically developed countries experienced lower mortality improvement rates after 2011, and this is consistent with the LC-GDP-CPI model.

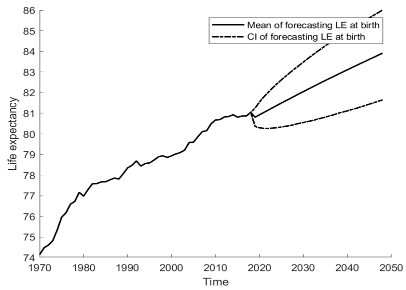
Figure 7 displays the period LE at birth in USA based on the four mortality models. The solid line is the expected period LE and black dash line is the 95%-CI. The LC model, LC-GDP model and LC-CPI model all yield an increasing trend, and the period life expectancies reach around 84 at the end of forecasting window. The LC-GDP-CPI model shows however a less optimistic period LE than other three models, with less than 82 years in the year 2048, and this is consistent with the findings of Figure 6.



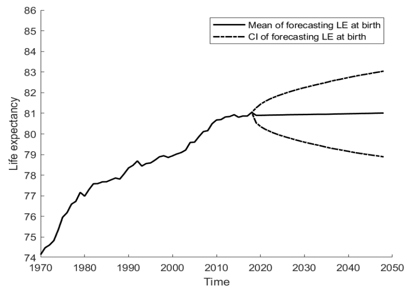
(a) LC



(b) LC-GDP



(c) LC-CPI



(d) LC-GDP-CPI

Figure 7: Period life expectancy (mean and 95%-confidence interval) in years for the USA based on the four mortality models.

6 Conclusion

This paper studies the impact of CPI on mortality dynamics, and how CPI can help to better predict future mortality. We use the log CPI as factor in mortality models, as the CPI closely approximates the costs of living for individuals. Our findings for the United States, Canada, Australia and France show that both the in-sample and out-of-sample performance is better if one replaces the log GDP in the model of Niu and Melenberg (2014) with log CPI.

The in-sample and out-of-sample performance are best for the mortality model that includes both the log GDP and log CPI factors, and this is evidence that economic growth helps us to model future trends in mortality. The in-sample performance of this model is also better than the Lee-Carter model, which uses one optimal but latent time trend. It is thus better to use the two observable time trends of log GDP and log CPI.

In conclusion, this paper proposes a new factor of economic growth to increase the accuracy of forecasting mortality. For further research, we suggest a more granular approach to study the impact of the standard of living on mortality dynamics.

Acknowledgement

We are grateful to the co-editor and two anonymous reviewers for comments and suggestions.

References

Agbonlahor, O. (2014). The impact of monetary policy on the economy of the United Kingdom: A vector error correction model (VECM). *European Scientific Journal 10*(16), 19–42.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica 66*(1), 47–78.

Bai, J. and P. Perron (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics 18*(1), 1–22.

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

Booth, H., J. Maindonald, and L. Smith (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies 56*(3), 325–336.

Bozzo, G., S. Levantesi, and M. Menzietti (2021). Longevity risk and economic growth in sub-populations: evidence from Italy. *Decisions in Economics and Finance 44*(1), 101–115.

Cupido, K., P. Jevtić, and T. J. Boonen (2020). Space, mortality, and economic growth. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3496270`.

De Jong, P. and L. Tickle (2006). Extending Lee–Carter mortality forecasting. *Mathematical Population Studies 13*(1), 1–18.

Djeundje, V. B., S. Haberman, M. Bajekal, and J. Lu (2022). The slowdown in mortality improvement rates 2011–2017: a multi-country analysis. *European Actuarial Journal 12*(2), 839–878.

Hanewald, K. (2011). Explaining mortality dynamics: The role of macroeconomic fluctuations and cause of death trends. *North American Actuarial Journal 15*(2), 290–314.

Kuang, D., B. Nielsen, and J. P. Nielsen (2008a). Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika 95*(4), 987–991.

Kuang, D., B. Nielsen, and J. P. Nielsen (2008b). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika 95*(4), 979–986.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics 54*(1-3), 159–178.

Lee, R. and T. Miller (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography 38*(4), 537–549.

Lee, R. D. and L. R. Carter (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association 87*(419), 659–671.

Leng, X. and L. Peng (2016). Inference pitfalls in Lee–Carter model for forecasting mortality. *Insurance: Mathematics and Economics 70*, 58–65.

Li, H., A. De Waegenaere, and B. Melenberg (2015). The choice of sample size for mortality forecasting: A Bayesian learning approach. *Insurance: Mathematics and Economics 63*, 153–168.

Liu, Q., C. Ling, D. Li, and L. Peng (2019a). Bias-corrected inference for a modified Lee–Carter mortality model. *ASTIN Bulletin: The Journal of the IAA 49*(2), 433–455.

Liu, Q., C. Ling, and L. Peng (2019b). Statistical inference for Lee-Carter mortality model and corresponding forecasts. *North American Actuarial Journal 23*(3), 335–363.

Luo, W. and Y. Xie (2020). Economic growth, income inequality and life expectancy in China. *Social Science & Medicine 256*, 113046.

Lyu, P., A. De Waegenaere, and B. Melenberg (2021). A multi-population approach to forecasting all-cause mortality using cause-of-death mortality data. *North American Actuarial Journal 25*(S1), S421–S456.

Nielsen, B. and J. P. Nielsen (2014). Identification and forecasting in mortality models. *The Scientific World Journal 2014*, 1–24.

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

O'Hare, C. and Y. Li (2012). Explaining young mortality. *Insurance: Mathematics and Economics 50*(1), 12–25.

Plat, R. (2009). Stochastic portfolio specific mortality and the quantification of mortality basis risk. *Insurance: Mathematics and Economics 45*(1), 123–132.

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

Seklecka, M., N. Md. Lazam, A. A. Pantelous, and C. O'Hare (2019). Mortality effects of economic fluctuations in selected eurozone countries. *Journal of Forecasting 38*(1), 39–62.

Shkolnikov, V. M., D. A. Jdanov, E. M. Andreev, and J. W. Vaupel (2011). Steep increase in best-practice cohort life expectancy. *Population and Development Review 37*(3), 419–434.

Van Berkum, F., K. Antonio, and M. Vellekoop (2016). The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal 2016*(7), 581–603.

Yang, Y. (2005). Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation. *Biometrika 92*(4), 937–950.

Yao, Y.-C. (1988). Estimating the number of change-points via Schwarz' criterion. *Statistics & Probability Letters 6*(3), 181–189.

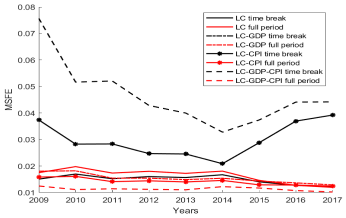
Zeileis, A., C. Kleiber, W. Krämer, and K. Hornik (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis 44*(1-2), 109–123.

Appendices

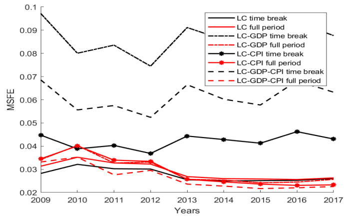
A. Structural break outcomes

Model	Countries	Number of break points	BIC	Final break-point
LC	USA	5	4.6942	1986
	Canada	5	38.0183	1987
	Australia	5	56.1860	1986
	France	5	37.4615	1995
LC-GDP	USA	4	-203.5872	2001
	Canada	4	-199.4683	2005
	Australia	4	-218.0377	1994
	France	4	-215.6357	1990
LC-CPI	USA	4	-227.1176	1997
	Canada	4	-250.8454	1991
	Australia	4	-232.3443	1991
	France	4	-251.9535	1992
LC-GDP-CPI	USA	1	-534.4544	1982
	Canada	1	-525.3156	1996
	Australia	1	-463.6942	2004
	France	1	-541.6455	1981

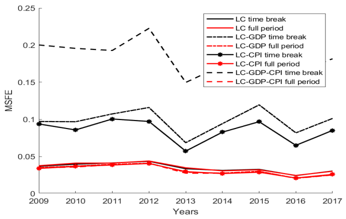
Table 13: The optimal number of break points with its BIC value and final break point year for all models and selected countries.



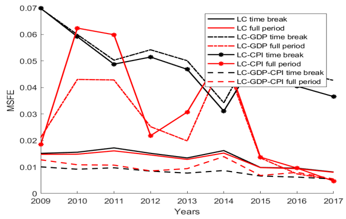
(a) The comparison of MSFE in USA based on the four mortality models with structural breaks.



(b) The comparison of MSFE in Canada based on the four mortality models with structural breaks.



(c) The comparison of MSFE in Australia based on the four mortality models with structural breaks.



(d) The comparison of MSFE in France based on the four mortality models with structural breaks.

Figure 8: The comparison of MSFE in the four selected countries and based on the four mortality models with structural breaks.