

Article

Markov-Switching Bayesian Vector Autoregression Model in Mortality Forecasting

Wanying Fu , Barry R. Smith , Patrick Brewer and Sean Droms 

Mathematics Department, Lebanon Valley College, Annville, PA 17003, USA; barsmith@lvc.edu (B.R.S.); brewer@lvc.edu (P.B.); droms@lvc.edu (S.D.)

* Correspondence: wfu@lvc.edu

Abstract: We apply a Markov-switching Bayesian vector autoregression (MSBVAR) model to mortality forecasting. MSBVAR has not previously been applied in this context, and our results show that it is a promising tool for mortality forecasting. Our model shows better forecasting accuracy than the Lee–Carter and Bayesian vector autoregressive (BVAR) models without regime-switching and while retaining the advantages of BVAR. MSBVAR provides more reliable estimates for parameter uncertainty and more flexibility in the shapes of point-forecast curves and shapes of confidence intervals than BVAR. Through regime-switching, MSBVAR helps to capture transitory changes in mortality and provides insightful quantitative information about mortality dynamics.

Keywords: MSBVAR; BVAR; regime-switching; mortality forecast; parameter projection; mortality structural change



Citation: Fu, Wanying, Barry R. Smith, Patrick Brewer, and Sean Droms. 2023. Markov-Switching Bayesian Vector Autoregression Model in Mortality Forecasting. *Risks* 11: 152. <https://doi.org/10.3390/risks11090152>

Academic Editor: Mogens Steffensen

Received: 30 June 2023

Revised: 17 August 2023

Accepted: 18 August 2023

Published: 22 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Accurate risk assessment and forecasting are essential for managing pensions, annuities, and other portfolios with mortality risk; yet, changing mortality patterns cause difficulties in mortality risk management and forecasting. Noting the importance of correlations between parameters and parameter uncertainty for mortality forecasting, [Njenga and Sherris \(2020\)](#) applied a Bayesian vector autoregressive (BVAR) model to forecast mortality. In this paper, we use a Markov-switching Bayesian autoregression model (MSBVAR) in mortality modeling for the first time. The resulting forecasts are more accurate than previous Bayesian-only forecasts, and have more reasonable confidence intervals. The model retains the strengths of BVAR, and is able to capture and quantify regime switches in mortality patterns, including shifts specific to particular age groups.

We follow a long-standing and frequently used forecasting methodology that begins with a parametric model for age-specific mortality. Many years of cross-sectional data are fit to this model and the parameters are modeled in turn as time series. In order to model the full age spectrum, many parameters are needed; one method of forecasting has been to forecast the time series for each parameter separately using univariate models. [Njenga and Sherris \(2020\)](#) applied multivariate models, namely, VAR and BVAR, to forecast the parameters in the [Heligman and Pollard \(1980\)](#) age-specific mortality model, which was the first time BVAR had been applied to a parametric model of the full age range. Using multivariate models is appropriate because there may be cross-correlation among parameter time series. They noted that BVAR is superior in several ways. BVAR provides more accurate forecasts than VAR, and cointegration is more easily handled; in addition, the Bayesian approach has a significant strength in that it can quantify both longevity risk and parameter uncertainty. However, the VAR and BVAR forecasts exhibit decreased accuracy compared to the well-known [Lee and Carter \(1992\)](#) model, and neither directly accounts for potential structural changes in mortality patterns, nor does the Lee–Carter model.

We adopt a similar methodology, except that we use a refined parametric model for the cross-sectional mortality and MSBVAR for the forecasts. Our MSBVAR forecasts (with

rolling cross-validation; see Section 3.3) are more accurate than Lee–Carter forecasts for one-, two-, and three-year forecasts. This accuracy persists, with average MSEs of one- to five-year forecasts that are over 30% lower than those produced by Lee–Carter. Moreover, the MSE of our MSBVAR model is less than half of that for the tested BVAR model with no regime switching, and our MSBVAR uses fewer parameters. For both BVAR and MSBVAR, we applied the reduced-form Sims–Zha Bayesian VAR model.

MSBVAR extends BVAR by including Markov-switching (MS) (also known as regime-switching behavior) to explicitly allow for structural changes in mortality patterns. The model automatically detects structural changes in which the training data appear to switch to a new regime, and uses BVAR as the model within each regime. The model assumes that regime switches follow a Markov process, and we assume for simplicity that there are only two regimes.

Using BVAR within each regime accounts for correlation between parameters, maintaining the advantages of BVAR identified by [Njenga and Sherris \(2020\)](#). For instance, Section 3.3.2 provides a three-year forecast and confidence interval for cross-sectional mortality of French civilian males. Direct comparison of MSBVAR and BVAR shows that MSBVAR provides more reasonably-behaved confidence intervals for forecast parameters. Consistently, MSBVAR provides narrower confidence intervals than BVAR for parameters that exhibit stable historical linear trends, although it provides wider confidence intervals for parameters that exhibit historical fluctuation or volatility. Moreover, point forecasts and confidence interval forecasts using MSBVAR exhibit greater flexibility in terms of their shape.

Other types of analysis show that MSBVAR is capturing two very different regimes. In Section 3.3.2, we provide three-year forecast densities for intercepts and AR(1) parameters when using MSBVAR with French civilian male data to forecast only those parameters associated with the youngest ages. Most of these distributions show a clear separation between the two regimes. The estimated error–covariance matrix within each regime reveals that one regime exhibits substantially lower volatility than the other (as in [Milidonis et al. \(2011\)](#)).

For our underlying age-specific parametric model, we selected an exemplar of the DLGC (differential logistic model with growth rate and controlling factor) framework in [Fu et al. \(2022\)](#), which models the full age cross-section of mortality (described in detail in Section 2). This work showed that such a model can provide a very close fit to historical data, and captures structural changes well enough that such data may become visible and salient features of the parameter time series plots. However, such identification by visual inspection is qualitative and subjective; one goal of using a regime-switching model in the present work is to provide a more objective determination of the timing and nature of mortality structural changes.

We can distinguish two types of changes: a once-and-for all permanent change to mortality patterns, for instance an improvement created by a major new medical technology, and transient changes that come and go. Only the latter may reasonably be described by a Markov process, although as noted, for instance in [Krolzig \(2000\)](#), “as the MS model also nests models with once-and-for-all structural breaks, it might be used to detect permanent breaks”. In this paper, we use the term “structural break” to refer to one-time permanent changes in mortality patterns, “regime switch” to refer to potentially random shifts that could be described by an MS model, and “structural change” as a generic term to encompass both types of change.

After fitting our age-specific model to French civilian male data from 1950–2020, we examined correlations among ten time series of parameters (after transforming them to be stationary). All except one of the ten transformed parameters naturally segregated into two groups that exhibited strong cross-correlation within each group and weak or no correlation between groups. These two groups of parameters conveniently corresponded to those associated with childhood and young adult mortality and those associated with middle-age and senescent mortality.

In Section 3.1, we fit univariate regime-switching models to each parameter time series in order to understand the nature of structural changes for individual parameters. We observed four parameters associated with adulthood mortality that underwent changes at around the same time in the early 1990s, while no such switching was present in the young-age or remaining old-age parameters. We then forecast the two groups of parameters separately with MSBVAR using a time interval for training that covered the early 1990s. The model again identified a regime switch in the group of older-age parameters around the 1990s and no such switch in the younger-age parameters. The common timing of all of these switches seemingly reflects changes in mortality in France amounting to an improvement during the 1990s, as depicted visually in parameter plots. Our model suggests that these changes in mortality only happened for adults, and did not impact childhood mortality.

The regime switching aspect of our model allows us to provide quantitative information for greater insight into structural mortality dynamics. The fitted transition matrix of the Markov process for adult and old-age mortality shows that there are 34.6 or 35.5 years (in each of the two regimes) on average between structural changes in adulthood and old-age mortality, while for infant and childhood mortality there are 65.8 or 68.5 years on average between structural changes in infant and childhood mortality.

The methodology described above, which fits a cross-sectional parametric model to mortality data and then forecasts parameter values, has long been used, for example, in McNown and Rogers (1989), Thompson et al. (1989), Avraam et al. (2014), Njenga and Sherris (2020), Fu et al. (2022), and Bardoutsos et al. (2018). The fitted parameters help to visualize mortality trends and reveal structural changes. One approach to mortality projection uses ARIMA models to forecast each parameter separately (see McNown and Rogers (1989), Fu et al. (2022), and Bardoutsos et al. (2018)). The forecast accuracy can depend on the age-specific parametric model used. A disadvantage of this approach is that it ignores correlations among the cross-sectional parameters of mortality, and may underestimate mortality risk.

Vector autoregression is commonly used as a multivariate forecasting model incorporating cross-correlation and cointegration. Njenga and Sherris (2020) applied both VAR and BVAR to mortality forecasting. Lu and Zhu (2023) recently applied a Bayesian version of VAR called FAVAR based on the Lee–Carter model, achieving somewhat better accuracy than Lee–Carter; however, this model uses over 6000 parameters, does not explicitly account for structural change, and has other drawbacks of Lee–Carter, such as a lack of natural interpretations for most parameters.

Njenga and Sherris (2020) suggested that their forecasts with VAR and BVAR (+Heligman–Pollard) were less accurate than Lee–Carter forecasts (see their Table 8) because their models employed far fewer parameters. An additional possibility is that structural changes in mortality patterns must be considered when projecting time series. Accounting for such changes could improve mortality forecasts, as demonstrated in Milidonis et al. (2011) and Fu et al. (2022). A number of works have dealt with observed structural changes by training the model only on data collected after the observed structural changes, i.e., structural changes are dealt with by limiting the training set. Lin and Liu (2007) created a different approach to describing age-specific mortality, using a Markov process to model “physiological age”, with the aim of tying the model more directly to biological mechanisms. As the authors noted, forecasting with such models may allow for the incorporation of expert opinions on medical or social impacts upon mortality, thereby integrating structural changes into forecasting.

By contrast, a regime-switching approach such as MSBVAR provides an automatic method to model the different mortality regimes. Regime-switching models have been applied to mortality modeling before, as in Gylys and Šiaulys (2020), Zhou (2019), Gao et al. (2015), Shen and Siu (2013), Ignatieva et al. (2016), Hainaut (2012) and Milidonis et al. (2011); the MSBVAR model has been used in econometric applications as well, as in Sims et al. (2008).

The remainder of this paper is organized as follows. Section 2 provides the model formulation, while Section 3 shows the main results. We begin in Section 3.1 with analysis that reveals and quantifies structural changes in individual parameters. After a brief

description of our MSBVAR forecasting strategy in Section 3.2, we present the results of accuracy testing of our model using rolling-window cross-validation and provide an example future mortality forecast in Section 3.3. Finally, Section 4 concludes the paper.

2. Model Formulation

In this paper, we apply an exemplar of the DLGC framework [Fu et al. \(2022\)](#) as the age-specific mortality model to estimate the yearly parameters and then forecast the estimated yearly parameters as time series by applying the MSBVAR model to the majority of the parameters and using means for the remaining.

2.1. Choosing $a(x)$ and $b(x)$ for the Age-Specific DLGC Model

The DLGC framework from [Fu et al. \(2022\)](#) that we use to model age-specific mortality provides strong fitting accuracy, out-performing the Heligman–Pollard model; it can reveal structural changes, and has good parameter interpretation. The strong fit promotes forecasting accuracy, the revelation of structural changes, some of which appear transitory, suggests the use of regime switching models, and the natural parameter interpretations provide insight into the structural changes.

The DLGC framework decomposes mortality into the three parts: infancy and childhood, adolescence and early adulthood (younger than age 30), and later adulthood to old age. Let $q(x)$ be the one-year mortality probability for a life aged x . The general model formula of DLGC for $q(x)$ is

$$q(x) = [1 + A^{-(B+x)^C}]^{-1} + \frac{T_m}{1 + \exp(-k(x-z))} + \frac{g \cdot \exp(\bar{b}(x)(x - M_1))}{\int_{M_1}^x \exp(\bar{b}(u)(u - M_1)) \cdot b(u)a(u)du + 2}, \quad (1)$$

where the three terms on the right side of the equation are the respective contributions to mortality of factors arising during the youngest ages, adolescence to early adulthood, and later adulthood to old age, denoted as $q_I(x)$, $q_T(x)$, and $q_{AO}(x)$, respectively. The model has eight constant parameters and two functions $a(x)$ and $b(x)$ (for later adulthood, including old age), chosen for convenience and need. In the third term, $\bar{b}(x)$ denotes the average value of $b(x)$ on $[M_1, x]$, i.e., $\bar{b}(x) = \int_{M_1}^x b(u)du / (x - M_1)$. The DLGC framework captures several human mortality features through its parameters, e.g., infant mortality (parameter A), the “accident hump” age in teenage years (parameter z), human asymptotic mortality ($T_m + g$), etc. [Table 1](#) provides more details and meanings for the parameters.

We carefully chose the functions $a(x)$ and $b(x)$ from the later adulthood part of our model. The function $a(x)$ has a range $[1, \infty)$ and represents the effects of beneficial factors, e.g., medical and scientific improvement, that tend to decelerate mortality probabilities. The function $b(x)$ increases with range $(0, 1)$ and represents the effect of natural factors, e.g., aging, that cumulatively accelerate mortality probability during later adulthood.

We define $a(x)$ as

$$a(x) = \frac{a_1 - 1}{1 + \exp(a_2(x - M_3))} + 1 \quad (2)$$

In this formula, $a(x)$ decreases from the maximum value a_1 to 1, reflecting our assumption that the decelerating effects of beneficial factors fade with age. The parameter M_3 is the age when $a(x)$ decreases the fastest, while a_2 models how fast $a(x)$ drops near age M_3 . The definition of $a(x)$ in [Fu et al. \(2022\)](#) partitions later adulthood ages into those less than 65, those between 65 and 85, and those greater than 85, and provides only the average effects of decelerating factors for the three subgroups. The above definition of $a(x)$ allows the model to capture the moving age and moving speed of the fastest decrease in decelerating effects.

Table 1. Parameters and functions assigned in the general model of $q(x)$.

	Pars	Meaning	Range	Other Constraints
Youngest age	A	the level of child mortality, approximate to $q(1)$	$(0, 1)$	
	B	the mortality probabilities' difference between age 0 and 1, i.e., $q(0) - q(1)$	$(0, 1)$	
	C	the decline in mortality during childhood	$(0, 1)$	
Teenage years	k	the growth rate of death probability only caused by teenage years factors	>0	
	z	the "accident hump" age	>0	$q_T(z) = \frac{1}{2}T_m$
	T_m	the maximum death probability only caused by teenage years factors	$(0, 1)$	
Later adulthood	M_1	the main increasing process of $q_{AO}(x)$ happens near M_1	>0	$q_{AO}(M_1) = \frac{1}{2}g$
	g	the maximum death probability only caused by later adulthood factors	$(0, 1)$	
	$b(x)$	growth rate of $q_{AO}(x)$ considering only natural and basic factors	$(0, 1)$	continuously increasing
	$a(x)$	a controlling factor on the increasing of death probability	$[1, \infty)$	continuous; satisfying $a(x) = 1$, for x after a certain old age

For the mortality accelerating factor $b(x)$, we keep the increasing logistic function used in the exemplar in [Fu et al. \(2022\)](#):

$$b(x) = \beta_1 + (\beta_3 - \beta_1) \cdot [1 + \exp(-\beta_2(x - M_2))]^{-1}. \tag{3}$$

The function $b(x)$ increases from its minimum value β_1 for young adulthood to its maximum value β_3 for very old age. It increases the fastest near age M_2 ; β_2 controls the increasing speed.

Our fitting results for male and female French civilians show that our age-specific model can capture both permanent and transitory structural changes in historical mortality and that the chosen forms of $a(x)$ and $b(x)$ successfully capture mortality structural changes in later adulthood (see details in Section 3.1).

2.2. MSBVAR

In this section, we apply the MSBVAR model described in [Sims et al. \(2008\)](#) to multiple time-series, i.e., mortality parameters, subject to regime switching.

We apply a Markov-switching framework in our model to capture those changes in mortality patterns that could reasonably be assumed to follow a Markov process. Assume that the system has h regimes, and let $H = \{1, 2, \dots, h\}$ denote them; then, $s_t \in H$ denotes the regime's state at time t . Let $q_{ij} = P(s_t = i | s_{t-1} = j)$ be the probability of s_t being i given that s_{t-1} is j . Then, $Q = (q_{ij})_{h \times h} \in [0, 1]^{h^2}$ is the Markov transition matrix and satisfies $\sum_{i \in H} q_{ij} = 1$.

Continuing with definitions, let $y_t \in \mathbb{R}^n$ denote the endogenous variables at time $t \geq 0$ and let $z_t \in \mathbb{R}^m$ be the predetermined exogenous variables (including an intercept) at time t , where we assume that y_t conditional on the past only depends on the current state, not on its lagged states. The MSBVAR model in [Sims et al. \(2008\)](#) considers a class of models for y_t and z_t (structural VAR with lag p) described below: given the initial conditions y_0, \dots, y_{1-p} ,

$$y_t' A(s_t) = \sum_{i=1}^p y_{t-i}' A_i(s_t) + z_t' C(s_t) + \epsilon_t' \Sigma^{-1}(s_t), 1 \leq t \leq T, \tag{4}$$

where:

- the prime notation represents the matrix transpose
- p is the lag length
- y_t is an n -dimensional column vector of endogenous variables at time t
- z_t is an m -dimensional column vector of exogenous and deterministic variables at time t
- ϵ_t is an n -dimensional column vector of unobserved random shocks at time t
- $A(k)$ is an invertible $n \times n$ matrix and $A_j(k)$ is an $n \times n$ matrix for $1 \leq k \leq h$
- $C(k)$ is an $m \times n$ matrix for $1 \leq k \leq h$
- $\Sigma(k)$ is an $n \times n$ diagonal matrix for $1 \leq k \leq h$, and ϵ_t , conditional on information until time $t - 1$, is assumed to follow the multivariate normal distribution with mean 0 and variance I_n , where I_n is the $n \times n$ identity matrix.

A VAR model requires estimates for all coefficients, while a BVAR model assumes that the coefficient matrices are random, assigning prior distributions and generating posterior distributions for coefficients. The well-known Litterman's Prior in Litterman (1986) and Robertson and Tallman (1999) specifies the prior distribution by first assigning them certain mean values and then measuring their variation (the distribution "tightness") around the given prior mean values based on a predetermined set of hyperparameters. Sims and Zha (1998) extended the Litterman prior as the Sims–Zha prior, and Sims et al. (2008) further combined the extended BVAR setting and the Markov-switching framework to introduce the MSBVAR model following Equation (4), which is the model implied in our mortality parameters projection. The set of hyperparameters used in MSBVAR contains the elements below (see Sims and Zha (1998) and Njenga and Sherris (2020)):

- $\lambda_0 \in [0, 1]$ is the overall tightness of the prior on the error covariance matrix. As it increases, the model moves away from a random walk.
- $\lambda_1 \in [0, 1]$ is the standard deviation or tightness of the prior around the AR(1) parameters.
- $\lambda_3 > 0$ is the lag decay; as it increases, it shrinks the higher-order lag coefficients to 0.
- $\lambda_4 > 0$ is the standard deviation or tightness around the intercept.
- $\lambda_5 > 0$ is a single value for the standard deviation or tightness around the exogenous variable coefficients.
- $\mu_5 \geq 0$ is the sum of the prior weights of the coefficients; larger values imply differences in stationarity.
- $\mu_6 \geq 0$ are dummy initial observations or prior drift; larger values allow for common trends.

In our mortality forecasting model in this paper, we apply and analyze the simplified MSBVAR case: an intercept is considered, with no other exogenous factors, in a one-lag ($p = 1$) and two-regime system ($h = 2$). Thus, the multiple time series y_t is only cross-correlated with its immediately previous value, and the model formulation in Equation (4) is simplified to

$$y_t' A(s_t) = \delta' + y_{t-1}' A_1(s_t) + \epsilon_t' \Sigma^{-1}(s_t), 1 \leq t \leq T. \quad (5)$$

where δ is an n -dimensional column vector of n constants.

In our approach, a Gibbs sampler is implemented to draw Bayesian posterior samples for the MSBVAR model, and we generate drawings from the posterior forecast density to provide the forecast and parameter uncertainty.

3. Main Results

In this section, we apply our model to the French civilian male population by age and year (1×1) using mortality data from the Human Mortality Database Human Mortality Database (n.d.) for the period of 1950–2020. We fit an age-specific model using the DLGC framework, with $a(x)$ and $b(x)$ chosen as described in Section 2.1 (denoted as DLGCab); Section 3.1 shows how the estimated parameters capture structural changes in mortality. In Section 3.2, we describe the MSBVAR forecasting strategy that we use to project the parameter time series in our paper. In Section 3.3, we perform an out-of-sample forecasting

test on our model using a form of rolling-window cross-validation, then we forecast the following three-year mortality of French male civilians after 2020.

3.1. Structural Breaks and Regime-Switching in the Fitted Parameters

In this section, we reveal mortality structural changes using plots of a selection of yearly DLGCab fitted parameters and supporting regime-switching models. While regime-switching models do not describe permanent one-time structural breaks, they can be used to detect both structural breaks and regime-switches, as observed by Krolzig (2000).

3.1.1. Mortality Structural Change during 1990s

For French male civilians (see Figure 1), $\log T_m$ shows a slope decrease during the 1990s, a_1 has an obvious increase starting in the late 1990s, M_2 shows a sudden boost in 1990, and M_3 shows an oscillating up-and-down trend including a “turning point” during the 1990s; the other parameters retain their existing trends or patterns through the 1990s. The trend changes in the four parameter plots imply that a structural change in mortality occurred at some point during the 1990s. Figure 2 shows that the corresponding plots for French female civilians have patterns similar to those for males for the period during and around the 1990s.

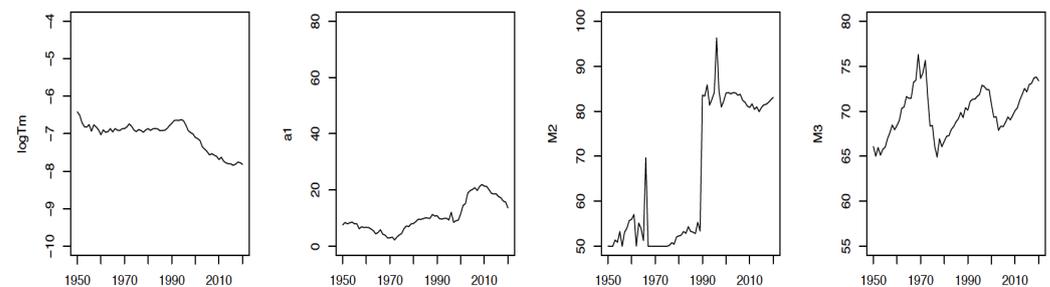


Figure 1. Fitted $\log T_m$, a_1 , M_2 , and M_3 for French male civilians during 1950–2020.

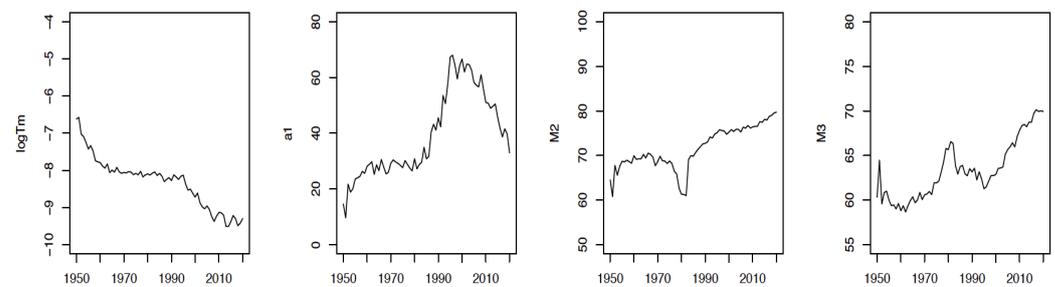


Figure 2. Fitted $\log T_m$, a_1 , M_2 , and M_3 for French female civilians during 1950–2020.

Three of the above four parameters are from the later adulthood portion of our model, and the other parameter is the asymptotic mortality from ages 14–30, which contributes to mortality in later adulthood. Thus, our model attributes the structural changes revealed in 1990s to changes in later adulthood mortality.

To quantify the notion that the trend changes shown by the parameter plots reflect structural changes in mortality, we modeled each parameter with a simple two-regime system by choosing a low lag AR model with small AIC in each regime. We did not try to find the best fit or to choose the best model; rather, the goal was to show that the trend changes described in the parameter plots can be identified mathematically and interpreted as structure changes mortality. The results reveal clear regime changes and two structural breaks, providing insightful quantified information about mortality structural changes based on the interpretation of the parameters.

A. M_2 Structural Change

We fit M_2 as a constant in each regime. Table 2 below shows the fitting results for regime-switching and the transition probability matrix for M_2 .

Table 2. Regime-switching model and transition matrix fitted for M_2 .

Coefficients			Transition Probabilities
	<i>Intercept(s)</i>	<i>Std(s)</i>	$Q = \begin{bmatrix} 0.9375 & 0.0256 \\ 0.0625 & 0.9744 \end{bmatrix}$
<i>Model 1</i>	82.5964	3.6150	
<i>Model 2</i>	52.0209	2.0645	

Figure 3 shows that a clear structural break happened in M_2 between 1990 and 1991. To interpret this change, recall that in our model $b(x)$ is the mortality accelerating factor, mainly expressing the aging effect, and M_2 is the age at which $b(x)$ increases the fastest.

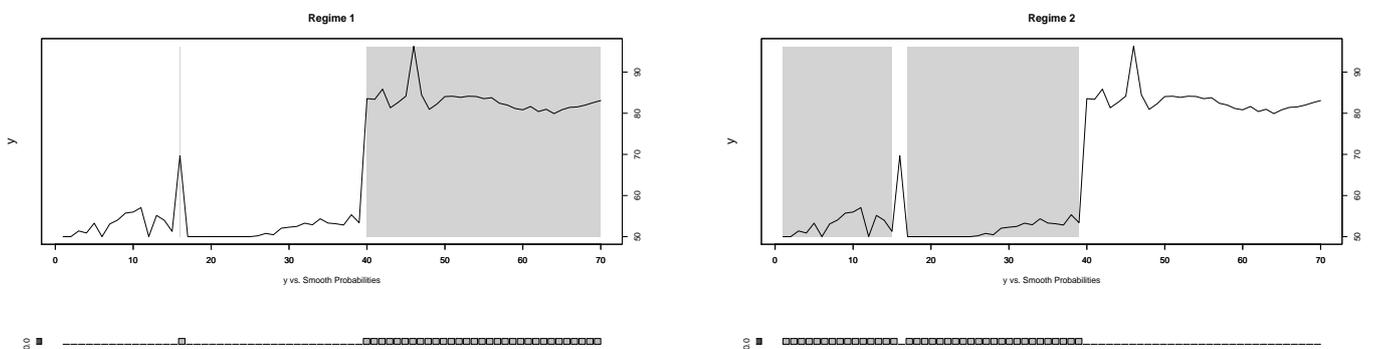


Figure 3. Regimes of fitted M_2 for French male civilians during 1950–2020. The horizontal scale is the number of years since 1950. The gray shading in each figure shows the regime indicated above the figure. The color switch at 40 represents the structural break between 1990 and 1991.

From the above fitted coefficients, it can be noticed that during 1950–1990, with the exception of 1967, the aging accelerative effect on mortality for French civilian males increases the fastest at about age 52.0209, with a volatility of 2.0645, while after 1991 the fastest increase is at about age 82.5964, with volatility 3.6150. This structural break shows a large improvement in in later adulthood mortality for French males during the 1990s.

From the plots and the transition matrix, we surmise that M_2 may have undergone one structural break around 1990 as well as transitory regime switches at other times.

B. $\log(T_m)$ Structural Change

We fit $\log(T_m)$ using an AR(1) model with an intercept in each regime. Table 3 below shows the fitting results for regime-switching and the transition probability matrix for $\log(T_m)$.

Table 3. Regime-switching model and transition matrix fitted for $\log(T_m)$.

Coefficients				Transition Probabilities
	<i>Intercept(s)</i>	<i>y₁(s)</i>	<i>Std(s)</i>	$Q = \begin{bmatrix} 1.0000 & 0.0406 \\ 5.3854 \times 10^{-9} & 0.9594 \end{bmatrix}$
<i>Model 1</i>	−2.0808	0.6963	0.0669	
<i>Model 2</i>	−0.7109	0.9106	0.0454	

Figure 4 shows a clear structural break in $\log(T_m)$ just before the 1990s. To interpret this fact, recall that T_m is the highest mortality accumulated by factors that occurred during the teenage years or young adulthood, i.e., asymptotic teenage and young adulthood mortality.

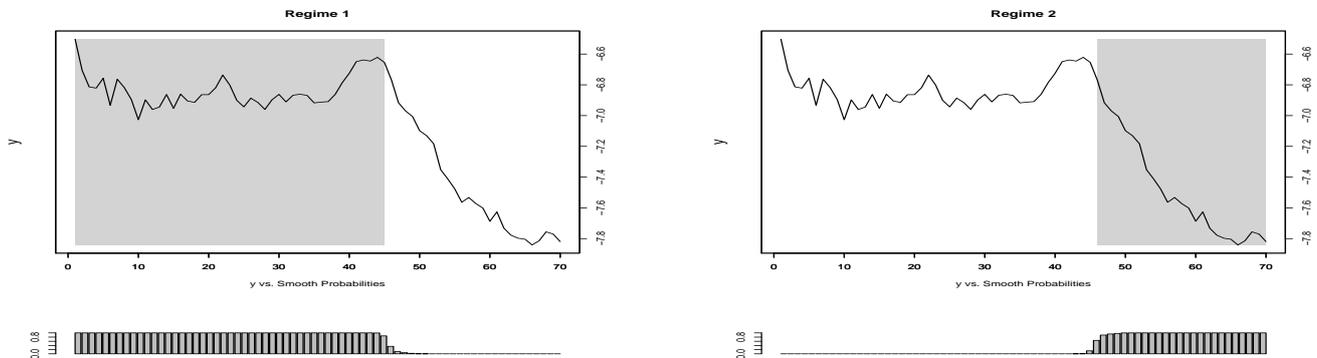


Figure 4. Regimes of fitted $\log(T_m)$ for French male civilians during 1950–2020. The horizontal scale is the number of years since 1950.

From the above fitted coefficients, it can be noticed that during 1950–1985 French male civilian asymptotic teenage and young adulthood mortality has a nearly flat pattern that is partially influenced by the previous year’s value, i.e., $\log T_m(t) = -2.0808 + 0.6963 \log T_m(t - 1)$. After 1986, $\log(T_m)$ has a declining pattern and follows $\log T_m(t) = -0.7109 + 0.9106 \log T_m(t - 1)$. This structural break shows a large improvement in mortality during adolescence and young adulthood for French males that occurred just before the 1990s. The transition matrix indicates that there is essentially zero probability of switching out of the current regime. This is consistent with $\log T_m$ having undergone no structural changes outside of the permanent break, as opposed to M_2 , which seems to have undergone both a permanent break and transitory changes.

C. a_1 Structural Change

We fit a_1 using an AR(1) model without an intercept in each regime. Table 4 below shows the fitting results for regime-switching and the transition probability matrix for a_1 .

Table 4. Regime-switching model and transition matrix fitted for a_1 .

Coefficients			Transition Probabilities
	$y_1(s)$	$Std(s)$	$Q = \begin{bmatrix} 0.8766 & 0.0215 \\ 0.1234 & 0.9785 \end{bmatrix}$
Model 1	1.0698	1.9740	
Model 2	0.9872	0.7871	

Figure 5 shows what can reasonably be interpreted as a regime switch in a_1 just before the 1990s (specifically, in 1986), followed by a switch back in the early 1990s. To interpret this fact, recall that $a(x)$ is the decelerative factor for later adulthood mortality and that a_1 is its highest value. In addition to the regime switch, the time plot itself shows a significantly different range of a_1 values before 1990 and after 1990.

In regime 1, $a_1(t) = 1.0698a_1(t - 1)$ shows a rapid increase, while in regime 2 $a_1(t) = 0.9872a_1(t - 1)$. From the transition matrix, regime 1 has a persistence of 0.8766 and jumping probability of 0.1234 in each year, which implies one jump every 8.1 years from regime 1 into regime 2. Regime 2 has a persistence of 0.9785 and jumping probability of 0.0215 in each year, which implies one jump every 46.5 years from regime 2 into regime 1.

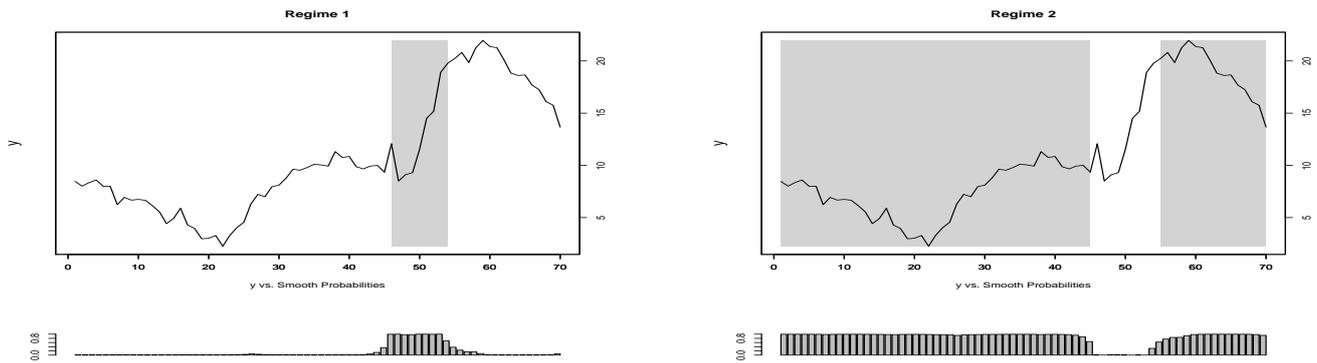


Figure 5. Regimes of fitted a_1 for French male civilians during 1950–2020. The horizontal scale is the number of years since 1950.

D. M_3 Oscillatory Pattern

We fit M_3 using an AR(2) model without an intercept in each regime. Table 5 shows the fitting results for regime-switching and the transition probability matrix for M_3 .

Table 5. Regime-switching model and transition matrix fitted for M_3 .

Coefficients				Transition Probabilities
	$y_1(s)$	$y_2(s)$	$Std(s)$	$Q = \begin{bmatrix} 0.9574 & 0.1257 \\ 0.0426 & 0.8743 \end{bmatrix}$
Model 1	0.5580	0.4495	0.4620	
Model 2	0.9407	0.0505	1.7721	

Figure 6 shows clear oscillatory structural changes in M_3 . To interpret this fact, recall that $a(x)$ is the decelerative factor for later adulthood mortality, and that it has its fastest period of declining at or near age M_3 .

The value of M_3 oscillates between the two regimes. From above fitted coefficients and transition matrix, in Regime 1, M_3 has an increasing pattern following the AR(2) process $M_3(t) = 0.5580M_3(t - 1) + 0.4495M_3(t - 2)$; M_3 has a persistence of 0.9574 in Regime 1 and a jumping probability of 0.0426 in each year, which implies one jump every 23.5 years from Regime 1 into Regime 2. In Regime 2, M_3 has a decreasing pattern following the AR(2) process $M_3(t) = 0.9407M_3(t - 1) + 0.0505M_3(t - 2)$; M_3 has a persistence of 0.8743 in Regime 2 and a jumping probability of 0.1257 in each year, which implies one jump every 8 years from Regime 2 into Regime 1.

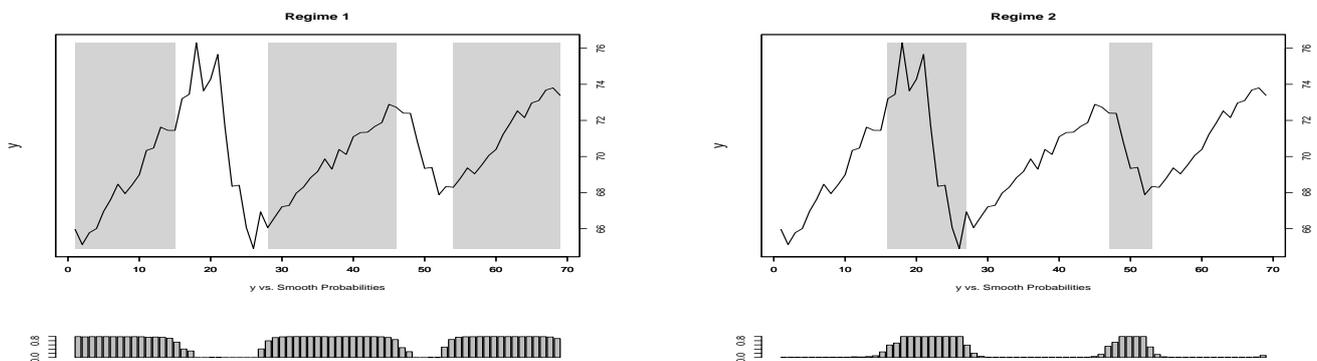


Figure 6. Regimes of fitted M_3 for French male civilians during 1950–2020. The horizontal scale is the number of years since 1950.

3.1.2. Structural Change in Mortality around the 1950s

Looking farther back in time, the age-specific model described in Section 2.1 additionally reveals structural changes in mortality in France during the 1950s. Figure 7 shows a sudden drop in $\log T_m$ near 1950 and a more negative slope for $\log A$ after 1950. No other parameters show structural changes at that time; thus, because T_m is the asymptotic mortality from the teenage years and A measures infant mortality, we can conclude that the structural change near 1950 resulted mainly from infant and teenage mortality. As analyzed in Fu et al. (2022), reasons for structural changes in mortality around the 1950s may include the wide production of penicillin in the mid-1940s and the use of antibiotics after abortion procedures, which improved survival for young females and infants.

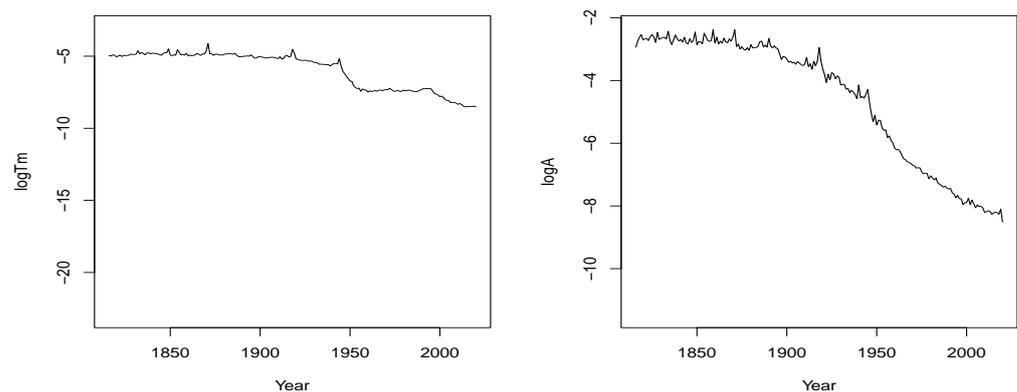


Figure 7. Fitted $\log T_m$ and $\log A$ for total French civilian population during the period of 1816–2020.

3.2. Forecasting Strategy Used in This Paper

In Section 3.1, we revealed structural change in mortality in the population of French male civilians during 1950–2020; in this section, we apply an MSBVAR model to forecast French civilian data and test its forecasting accuracy for the period of 1953–2020. Using the yearly fitted parameters of the age-specific DLGCab model described in Section 2.1, we design our forecasting strategy as follows: constant mean for parameters $k, z, a_2, \beta_1, \beta_2$, MSBVAR on the group of parameters A, B, C , and MSBVAR on the group of parameters $T_m, M_1, a_1, \beta_3, M_2, M_3, g$. Before applying MSBVAR, we transform the parameters to ensure that they are stationary; this process is described in Section 3.2.2.

3.2.1. Constant Mean

We set each of the five parameters $k, z, a_2, \beta_2, \beta_1$ equal to the average of their respective fitted yearly values because of the lack of any trend in their values, as well as to decrease the number of parameters in the forecast that need to be fit. More specifically:

- The values of a_2 and β_2 show high volatility and a noisy pattern; thus, a constant mean is a reasonable choice for forecasting them.
- The values of β_1, k and z are stable near a constant for a number of years; for example, β_1 very stably remains near 0.1 over several years, which matches the findings in de Beer and Janssen (2016) and Thatcher (1999) that from soon after 30 years of age the probability of dying increases by about 10% with each successive year of age.

3.2.2. Two Groups of Transformed Parameters to Apply MSBVAR

We apply log and order-1 or order-2 difference on each of the remaining ten parameters (not including the five parameters fitted by the constant mean) to make each time series closer to stationary. After these transformations, all ten transformed parameters pass the ADF and KPSS tests for stationarity.

The correlation matrix of the transformed parameters is shown in Table 6. The transformations applied to each parameter are shown by the row/column names of the matrix, where “log” means that we take the log of the parameter first, “diff” means that we take the

order-1 difference after the log, and “diff2” means that we take the order-2 difference after the log. For example, “log M_3_diff2 ” indicates the values of twice-differenced $\log(M_3)$.

The values in the correlation matrix suggest dividing the parameters into two groups and applying MSBVAR to each group.

- Group 1 (A, B, C): All pairwise correlations are larger than 0.831, with the highest value being 0.959; furthermore, there is no correlation larger than 0.235 in absolute value between the parameters in group 1 and the parameters outside this group, and only those with the single parameter M_2 are above 0.1.
- Group 2 ($T_m, M_1, a_1, \beta_3, M_2, M_3, g$): Several of the correlations in this group have a magnitude of at least 0.349, with three being greater than 0.87, and most of the seven parameters are included in these correlations.

The above grouping of the parameters is supported by their interpretations and usage in our model. The parameters in Group 1 are all young-age mortality parameters (i.e., in the youngest-age mortality q_I), while the parameters in Group 2 are all adult and old-age mortality parameters (i.e., in q_T and q_{AO}).

Table 6. Correlation matrix for the ten transformed parameters.

	logA_diff	logB_diff	logC_diff	logT _m _diff2	logM ₁ _diff	loga ₁ _diff2	logβ ₃ _diff	logM ₂ _diff	logM ₃ _diff2	logg_diff
logA_diff	1.000	0.846	0.831	0.016	0.040	−0.047	0.014	0.111	0.019	0.003
logB_diff	0.846	1.000	0.959	−0.002	0.065	0.018	−0.007	0.216	−0.004	0.003
logC_diff	0.831	0.959	1.000	0.038	0.089	0.058	−0.025	0.235	−0.006	0.017
logT _m _diff2	0.016	−0.002	0.038	1.000	0.222	−0.172	−0.123	0.289	0.020	−0.022
logM ₁ _diff	0.040	0.065	0.089	0.222	1.000	0.283	−0.934	0.225	0.099	0.872
loga ₁ _diff2	−0.047	0.018	0.058	−0.172	0.283	1.000	−0.349	0.274	−0.547	0.405
logβ ₃ _diff	0.014	−0.007	−0.025	−0.123	−0.934	−0.349	1.000	−0.230	−0.104	−0.927
logM ₂ _diff	0.111	0.216	0.235	0.289	0.225	0.274	−0.230	1.000	−0.174	0.118
logM ₃ _diff2	0.019	−0.004	−0.006	0.020	0.099	−0.547	−0.104	−0.174	1.000	0.083
logg_diff	0.003	0.003	0.017	−0.022	0.872	0.405	−0.927	0.118	0.083	1.000

Next, we describe the translation of our model into the abstract description of the MSBVAR model described above. From Equation (5), after post-multiplying through the invertible $A(s_t)^{-1}$, taking the Group 1 parameters as an example, in our approach the transformed young-age mortality parameters $y_t = (\log(A_t) - \log(A_{t-1}), \log(B_t) - \log(B_{t-1}), \log(C_t) - \log(C_{t-1}))'$ in the certain regime s_t (equaleither to 1 or 2) at time t satisfy the following equation:

$$y_t = c(s_t) + B'(s_t)y_{t-1} + e_t, \tag{6}$$

where $c(s_t) = (c_1(s_t), c_2(s_t), c_3(s_t))' = (A(s_t)^{-1})'\delta$ is the unknown intercept vector, $B(s_t) = A_1(s_t)A(s_t)^{-1}$ is an unknown 3×3 matrix of coefficients at lag 1, and e_t represents independent identically-distributed errors distributed as $N_3(0, S)$ with the unknown error covariance matrix $S = (A(s_t)A(s_t)')^{-1}$. Equation (6) can be written in matrix form as follows:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} c_{1t}(s_t) \\ c_{2t}(s_t) \\ c_{3t}(s_t) \end{pmatrix} + \begin{pmatrix} B_{11}(s_t) & B_{12}(s_t) & B_{13}(s_t) \\ B_{21}(s_t) & B_{22}(s_t) & B_{23}(s_t) \\ B_{31}(s_t) & B_{32}(s_t) & B_{33}(s_t) \end{pmatrix}' \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{pmatrix} \tag{7}$$

We then treat the unknown VAR parameters above in a Bayesian context as jointly distributed random variables by using the Sims–Zha normal–flat prior to fit and forecast the mortality parameters y_t with MSBVAR. We follow a similar process to forecast the mortality parameters in Group 2.

3.3. Forecasting Results

3.3.1. Out-of-Sample Forecasting Comparison Using Rolling-Window Cross-Validation

We applied rolling-window cross-validation to compare the out-of-sample forecasting accuracy of four models: (1) our model from Section 2.1 (DLGCab) combined with MSBVAR as described in Section 3.2; (2) the log-Poisson Lee–Carter model Brouhns et al. (2002); (3) VAR with lag-1 combined with our age-specific DLGCab model; and (4) BVAR with lag-1 combined with our age-specific DLGCab model. The VAR and BVAR models were both applied to twelve DLGCab parameters, while the remaining three, β_1 , β_2 , and a_2 , were set to be constant using their means. Therefore, there were comparable numbers of parameters when using the MSBVAR and BVAR models for multivariate time series forecasting (234 for BVAR and 208 for two-group MSBVAR). Moreover, we used the Sims–Zha normal–flat prior for both MSBVAR and BVAR, and used similar strategies to chose the hyperparameter values (see footnotes below for details). Training and testing were based on mortality data from French male civilians from 1953–2020 for ages 0 to 105; the testing results are shown in Table 7.

Table 7. One-year to three-year forecasting accuracy comparison of four models’ MSE of q_x for French male civilian population.

Training Period	Forecast Period	LC_logPoi	VAR(1)	BVAR(1)	Two-Group MSBVAR(1)
1953–1998	1999–2001	4.058080×10^{-5}	1.846671×10^{-5}	8.634617×10^{-5}	1.652785×10^{-5}
1954–1999	2000–2002	3.096344×10^{-5}	3.636803×10^{-6}	5.286439×10^{-5}	1.619272×10^{-5}
1955–2000	2001–2003	3.998948×10^{-5}	2.027850×10^{-5}	2.199262×10^{-5}	1.654713×10^{-5}
1956–2001	2002–2004	3.965374×10^{-5}	1.105280×10^{-4}	1.361299×10^{-4}	1.981227×10^{-5}
1957–2002	2003–2005	4.237491×10^{-5}	4.750569×10^{-5}	6.453116×10^{-5}	4.875723×10^{-5}
1958–2003	2004–2006	5.315747×10^{-5}	1.373889×10^{-3}	1.389112×10^{-4}	5.079026×10^{-5}
1959–2004	2005–2007	3.805303×10^{-5}	2.024754×10^{-2}	4.464989×10^{-5}	4.506862×10^{-5}
1960–2005	2006–2008	3.819256×10^{-5}	6.901466×10^{-6}	3.070929×10^{-5}	2.972596×10^{-5}
1961–2006	2007–2009	2.166763×10^{-5}	2.789243×10^{-5}	3.493766×10^{-5}	2.536762×10^{-5}
1962–2007	2008–2010	2.578418×10^{-5}	1.371940×10^{-5}	3.575074×10^{-5}	1.345337×10^{-5}
1963–2008	2009–2011	3.077442×10^{-5}	8.987640×10^{-6}	1.806542×10^{-4}	3.855534×10^{-6}
1964–2009	2010–2012	3.788919×10^{-5}	1.593804×10^{-5}	2.633839×10^{-5}	8.409379×10^{-6}
1965–2010	2011–2013	3.361986×10^{-5}	4.207817×10^{-5}	1.243555×10^{-5}	1.548707×10^{-5}
1966–2011	2012–2014	2.619416×10^{-5}	2.614010×10^{-5}	4.466430×10^{-5}	1.595633×10^{-5}
1967–2012	2013–2015	2.357353×10^{-5}	3.341852×10^{-5}	4.016176×10^{-5}	1.841944×10^{-5}
1968–2013	2014–2016	2.106681×10^{-5}	1.134660×10^{-5}	3.815115×10^{-5}	1.343410×10^{-5}
1969–2014	2015–2017	1.018989×10^{-5}	7.637350×10^{-6}	5.350254×10^{-5}	9.641061×10^{-5}
1970–2015	2016–2018	9.301274×10^{-6}	5.500549×10^{-6}	3.125023×10^{-5}	4.138879×10^{-5}
1971–2016	2017–2019	2.765426×10^{-5}	7.897331×10^{-6}	1.665924×10^{-5}	2.939270×10^{-6}
1972–2017	2018–2020	4.591307×10^{-5}	2.929639×10^{-5}	2.578495×10^{-5}	1.817020×10^{-5}

Each cell of Table 7 reports the average MSE of q_x for a one-, two-, and three-year forecast based on training on the previous 46 years’ data.^{1,2} For example, the second row of the table indicates that we trained the models on data from 1953–1998 and forecast values of q_x for 1999–2001. The subsequent four cells report the average MSE of the q_x estimates for the four models.

The sum of the MSEs of q_x for all sixty forecasts are shown below for each model; each sum equals the respective column sum multiplied by 3:

Model	LC_logPoi	VAR(1)	BVAR(1)	Two-group MSBVAR(1)
Total MSE of q_x	0.001909781	0.066175792	0.003349276	0.001550141

The two-group MSBVAR(1) has the lowest total MSE among the four models. Three of the models, Var(1), BVAR(1) and MSBVAR(1), consider the correlations among parameters when forecasting. By applying Markov switching, the two-group MSBVAR reduces the total MSE to about half that of the BVAR(1) forecast and about 1/40 of that of VAR(1), while also fitting the fewest parameters. This supports our conjecture in the introduction that mortality structural changes should be considered in mortality forecasting and that doing so can improve forecast accuracy over a BVAR model. The two-group MSBVAR(1) has a total MSE that is 18.83% less than that of the log-Poisson Lee–Carter model. Furthermore,

Section 3.3.2 shows that the two-group MSBVAR(1) provides more reasonable parameter confidence intervals than BVAR(1), and can provide insightful information about the structural changes and dynamics of intrinsic factors involved in mortality.

To broaden the scope of the forecasting accuracy, we carried out a one-year to five-year comparison of forecasting accuracy between our model and the log-Poisson Lee-Carter model. The result are recorded in Table 8, which reports the average testing MSE of the one-year to five-year q_x forecasts for the two models. The training and testing were based on French male civilian mortality data from 1962 to 2015 for ages 0 to 105, and rolling-window cross-validation was applied.

From the Table 8, it can be seen our model outperforms the log-Poisson Lee-Carter model in the five-year forecasting horizon. Based on the sum of the testing MSEs of q_x for all 45 forecasts, our model has a total MSE 32.14% less than that of the log-Poisson Lee-Carter model, and has eight lower average MSEs out of the nine one-year to five-year forecasting tests.

Table 8. One-year to five-year forecasting accuracy comparing four models’ MSE of q_x for the French male civilian population.

Training Period	Forecast Period	LC_logPoi	Two-Group MSBVAR(1)
1962–2007	2008–2012	1.122421×10^{-4}	2.328074×10^{-5}
1963–2008	2009–2013	1.263967×10^{-4}	1.305586×10^{-5}
1964–2009	2010–2014	1.207380×10^{-4}	8.582214×10^{-6}
1965–2010	2011–2015	1.001330×10^{-4}	2.192817×10^{-5}
1966–2011	2012–2016	1.409463×10^{-4}	8.506634×10^{-5}
1967–2012	2013–2017	1.129325×10^{-4}	2.735066×10^{-5}
1968–2013	2014–2018	5.705891×10^{-5}	1.345621×10^{-5}
1969–2014	2015–2019	1.030199×10^{-4}	1.351506×10^{-4}
1970–2015	2016–2020	1.015846×10^{-4}	6.914451×10^{-5}
Sum of the MSEs of q_x for all 45 forecasts		2.925156×10^{-3}	1.985076×10^{-3}

Njenga and Sherris (2020) used the Heligman–Pollard (HP) age-specific parametric mortality model to introduce BVAR forecasting. In this paper, we applied MSBVAR to a different age-specific mortality model, DLGC, and obtained better forecasting accuracy. To understand how much of this improvement is based on the age-specific model (HP versus DLGC) and how much is due to the forecasting method (BVAR versus MSBVAR), we conducted further experiments; specifically, we compared HP with BVAR, DLGC with BVAR, and DLGC with MSBVAR (i.e., HP+BVAR with models (1) and (4) mentioned at the beginning of this subsection) using the same methods and data as above. Through application of logarithm or first- or second-order differencing, six of the eight HP parameters were successfully transformed into stationary series, although two of them could not be transformed (as opposed to DLGC, for which all parameters successfully transformed to stationary). After examining the correlations among the transformed HP parameters, we observed no separation into groups, as was seen with DLGC; thus, we forecast all transformed HP parameters together as a group. We used rolling-window cross-validation, as in Table 7, except that we shifted the window four years each time, with just five training and forecast periods, and computed the sum of the MSEs for out-of-sample forecasts one, two, and three years ahead. In switching from HP+BVAR to DLGC+BVAR, the sum of the MSEs dropped by a factor of approximately 12, while when comparing DLGC+BVAR with DLGC+MSBVAR they dropped by a factor of about 3.5. Thus, the choice of DLGC as the base model seems to have had a significant impact on the improved forecasting accuracy achieved in the present study.

One potential reason for this is that HP has grown worse at fitting age-specific mortality data in recent decades. The blue curve in Figure 8 plots the MSE when fitting HP to one-year mortality data for French male civilians ages 0–105 against time. The red curve shows the MSE during the same years when using DLGC instead. The decrease in accuracy of HP over time is striking; a closer look at the data indicates that the decrease is largely due

to HP having a poor fit to old-age mortality. DLGC's continued accuracy is presumably because it has more parameters for modeling old-age mortality.

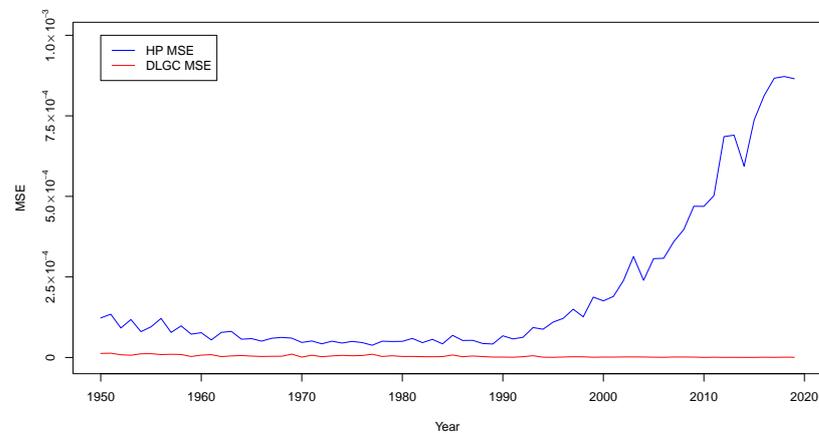


Figure 8. Comparing the fit of HP and DLGC to yearly mortality data from French male civilians.

3.3.2. Future Mortality Forecasting Example

We forecast the 2023 mortality q_x for age 0–105 for the French male civilian population by training on the period 1953–2020 using the same two-group MSBVAR(1) model as in Sections 3.2 and 3.3.1. The hyperparameter values were chosen as follows: for Group 1 transformed parameters fitted by MSBVAR with lag 1 and two regimes, $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_3 = 1$, $\lambda_4 = 0.1$, $\lambda_5 = 0$, $\mu_5 = 0$, $\mu_6 = 0$; for Group 2 transformed parameters fitted by MSBVAR with lag 1 and two regimes, $\lambda_0 = 1$, $\lambda_1 = 0.4$, $\lambda_3 = 1$, $\lambda_4 = 0.15$, $\lambda_5 = 0$, $\mu_5 = 0$, $\mu_6 = 0$. The prior for both was the Sims–Zha normal–flat prior.

Figure 9 shows the mortality forecasting results and the 68% and 90% confidence intervals.

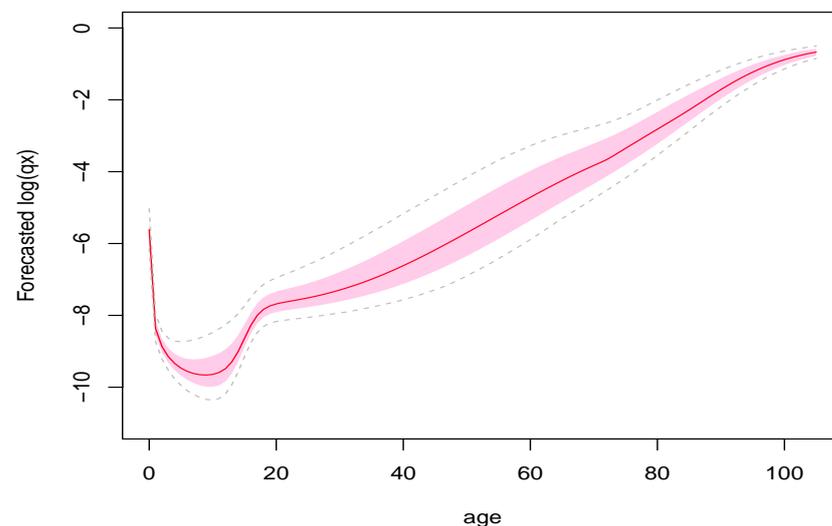


Figure 9. Fitted 2023 $\log q_x$ for French male civilian population with a training period of 1953–2020 for the DLGCab+two-group MSBVAR model for ages 0–105. The 68% confidence intervals of the forecasts are shown in pink, while the 90% confidence intervals are marked as grey dashed lines.

Our use of MSBVAR enables us to estimate the uncertainty in the parameters. Figure 10 provides the 68% confidence intervals for the one-year through ten-year forecasts of the ten time series parameters estimated by our model. Here, we include the Sims–Zha normal–flat BVAR(1), i.e., model (4) described in Section 3.3.1, as the comparison model. The estimates

come from 20,000 drawings from the associated posterior forecast density for both models. The comparison shows several advantages of MSBVAR over BVAR(1):

1. For parameters showing a stable historical trend, such as $\log A$, B , C , and β_3 (i.e., an almost linear trend), the MSBVAR model consistently provides much narrower 68% confidence intervals than the BVAR(1) model. Conversely, for parameters showing higher volatility in their historical values, such as M_1 , a_1 , M_2 , and M_3 , the MSBVAR model provides much wider 68% confidence interval than that BVAR(1). Thus, MSBVAR forecasts are more reasonable continuations of historically observed trends and volatility as compared with BVAR(1). A possible explanation for this advantage is that MSBVAR can recognize a period of time as being in a high-volatility or low-volatility regime, then use less/more volatile distribution estimates for the VAR coefficients when forecasting in the respective regimes. This explanation is further analyzed and supported later in this section.
2. As shown by the red/blue curve, i.e., the forecast value for each parameter, BVAR(1) tends to provide a forecast curve with a linear shape to describe only a coarse general tendency, while our model can provide a more flexibly shaped forecast curve when necessary by considering the historical regime switches in the parameter. For example, our model predicts a_1 will first drop and then increase in the following ten years, with a “hook” shaped curve, which reflects the historical ups and downs of a_1 . However, BVAR(1) only predicts the coarse increasing tendency with a line segment. Another example is M_3 ; BVAR(1) forecasts that it will continue the upward trend of the most recent historical data, while MSBVAR forecasts that it will decrease, beginning another one of the regularly occurring downward trends visible in the historical plot of M_3 . The pink confidence intervals are more flexible in their shapes, often with asymmetric behavior between the upper and lower bounds. Both a_1 and M_3 experience structural changes in the period 1953–2020 (see Section 3.1 for details).
3. Based on the parameter interpretation in Table 1, BVAR(1)'s one-year through ten-year confidence intervals include several unreasonable prediction values for a number of parameters, while those of our model provide predictions that are within range. Examples include: for BVAR(1), the confidence intervals for A (infant mortality) and B (decline in mortality at age 1) cover values very close to 1; C (childhood mortality decline) and g (the old-age component of asymptotic mortality) have confidence intervals that include values of 1 and even greater; and the confidence interval for β_3 (the cap on mortality growth rate for old age) includes negative values. The MSBVAR model provides more acceptable forecast confidence intervals for all of these parameters.

Below, we describe the two-regime MSBVAR(1) fitting results with respect to regime change and provide the transition matrix for the two groups of parameters, where Group 1 (i.e., A, B, C) includes the majority of parameters in childhood and teenage mortality of DLGCab and Group 2 (i.e., $T_m, M_1, a_1, \beta_3, M_2, M_3, g$) includes most of parameters in adulthood and old-age mortality of DLGCab.

For the Group 1 parameters, Figure 11 shows that infant and childhood mortality experience structural changes near 1967 and 2002. The Group 1 MSBVAR fitted full mean transition matrix in Table 9 shows an estimate for the infant and childhood mortality regime-switching transition matrix. From the transition matrix, regime 1 has persistence of 0.98540995. When in regime 1, the probability of a jump event into regime 2 in each year is 0.01459005; thus, we can expect about one jump per 68.5 years. Regime 2 has a persistence of 0.98479471; thus, when in regime 2, the probability of a jump event into regime 1 is 0.01520529 in each year, and we can expect about one jump per 65.8 years.

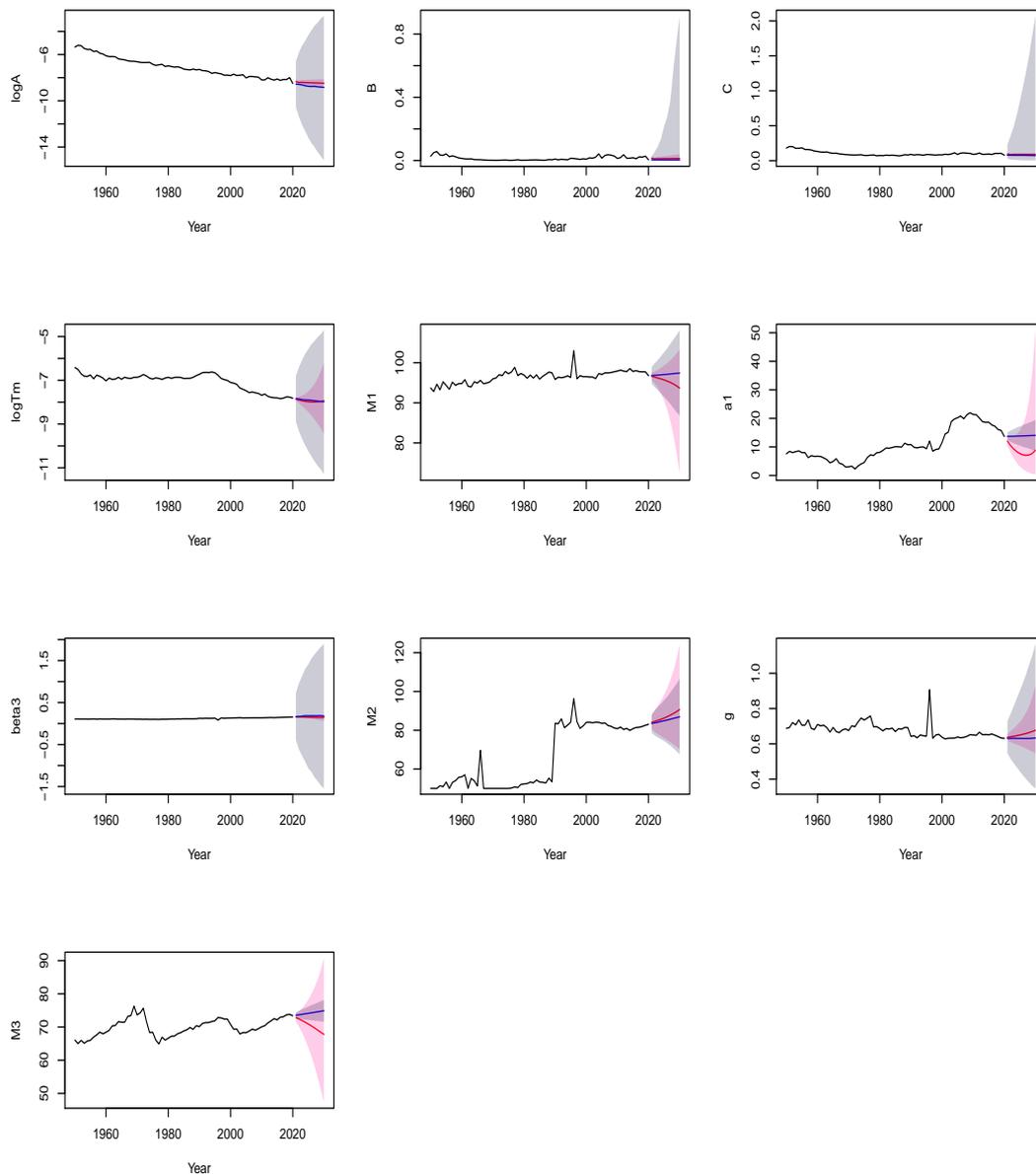


Figure 10. Time series of parameter values fit to historical data and 68% confidence intervals for one-year through ten-year forecasts with both two-group MSBVAR(1) and BVAR(1) after training on 1953–2020 French male civilian population data. In each plot, the solid black curve shows the estimated historical values for the parameter from fitting with DLGCab over 1953–2020, the red curve and solid blue curve represent ten years of point-forecast values from the two-group MSBVAR(1) and BVAR(1) models, respectively, and the pink shaded part and grey shaded part represent the forecast 68% confidence intervals from the two-group MSBVAR(1) and BVAR(1) models, respectively.

Table 9. Full mean transition matrix for the two groups of transformed parameters fitted by two-group MSBVAR(1).

Full Mean Transition Matrix for Group 1 Parameters	Full Mean Transition Matrix for Group 2 Parameters
$Q = \begin{bmatrix} 0.98540995 & 0.01520529 \\ 0.01459005 & 0.98479471 \end{bmatrix}$	$Q = \begin{bmatrix} 0.97106085 & 0.02810913 \\ 0.02893915 & 0.97189087 \end{bmatrix}$

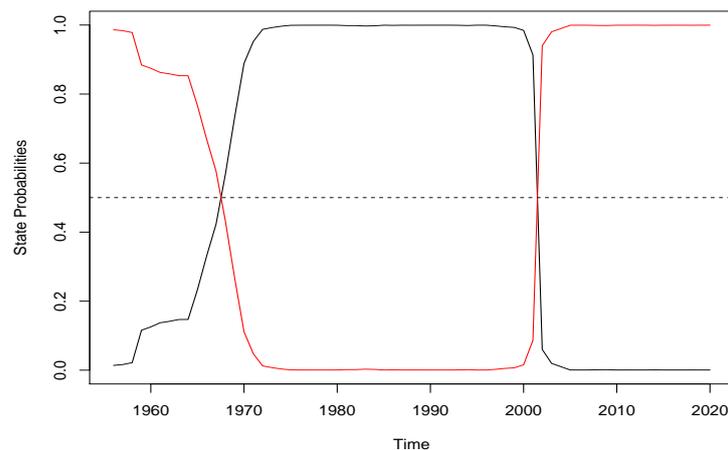


Figure 11. Fitted state probabilities from two-regime MSBVAR(1) for the Group 1 transformed parameters during 1953–2020. The black curve expresses the state probability of regime 1 while the red curve expresses the state probability of regime 2.

For the Group 2 parameters, Figure 12 shows that later adulthood mortality experiences oscillatory structural changes and alternates between the two regimes. This result matches our finding about later adulthood mortality structural changes during the 1990s and the individual parameter regime switching analysis (e.g., an oscillatory pattern in M_3) in Section 3.1.

The Group 2 MSBVAR fitted full mean transition matrix in Table 9 shows an estimate for the later adulthood regime-switching transition matrix. From the transition matrix, regime 1 has a persistence of 0.97106085. When in regime 1, the probability of a jump event into regime 2 each year is 0.02893915; thus, we can expect about one jump per 34.6 years. Regime 2 has a persistence of 0.97189087; thus, when in regime 2 the probability of a jump event into regime 1 is 0.02810913 each year, and we can expect about one jump per 35.6 years.

Figure 13 shows the intercepts and AR(1) coefficient posterior densities by regime for the Group 1 parameters (transformed and scaled by 100), which come from 20,000 drawings from the posterior forecast density in our forecasts. As can be seen, the distributions of most parameters show a clear separation between the two regimes. Many of the transformed Group 2 parameters show clear separation as well, although we have omitted the plots here.

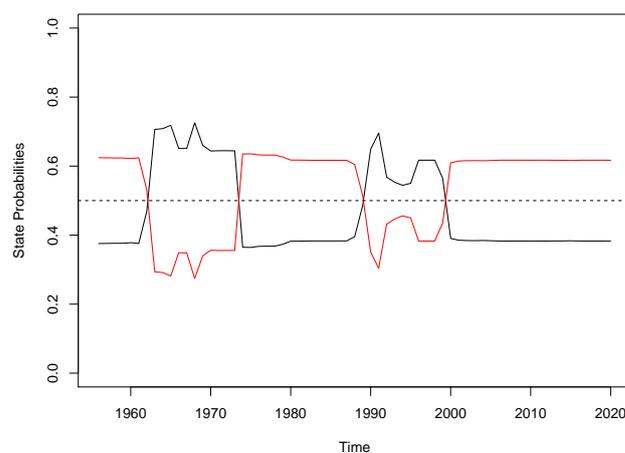


Figure 12. Fitted state probabilities from two-regime MSBVAR(1) for the Group 2 transformed parameters during 1953–2020. The black curve expresses the state probability of regime 1, while the red curve expresses the state probability of regime 2.

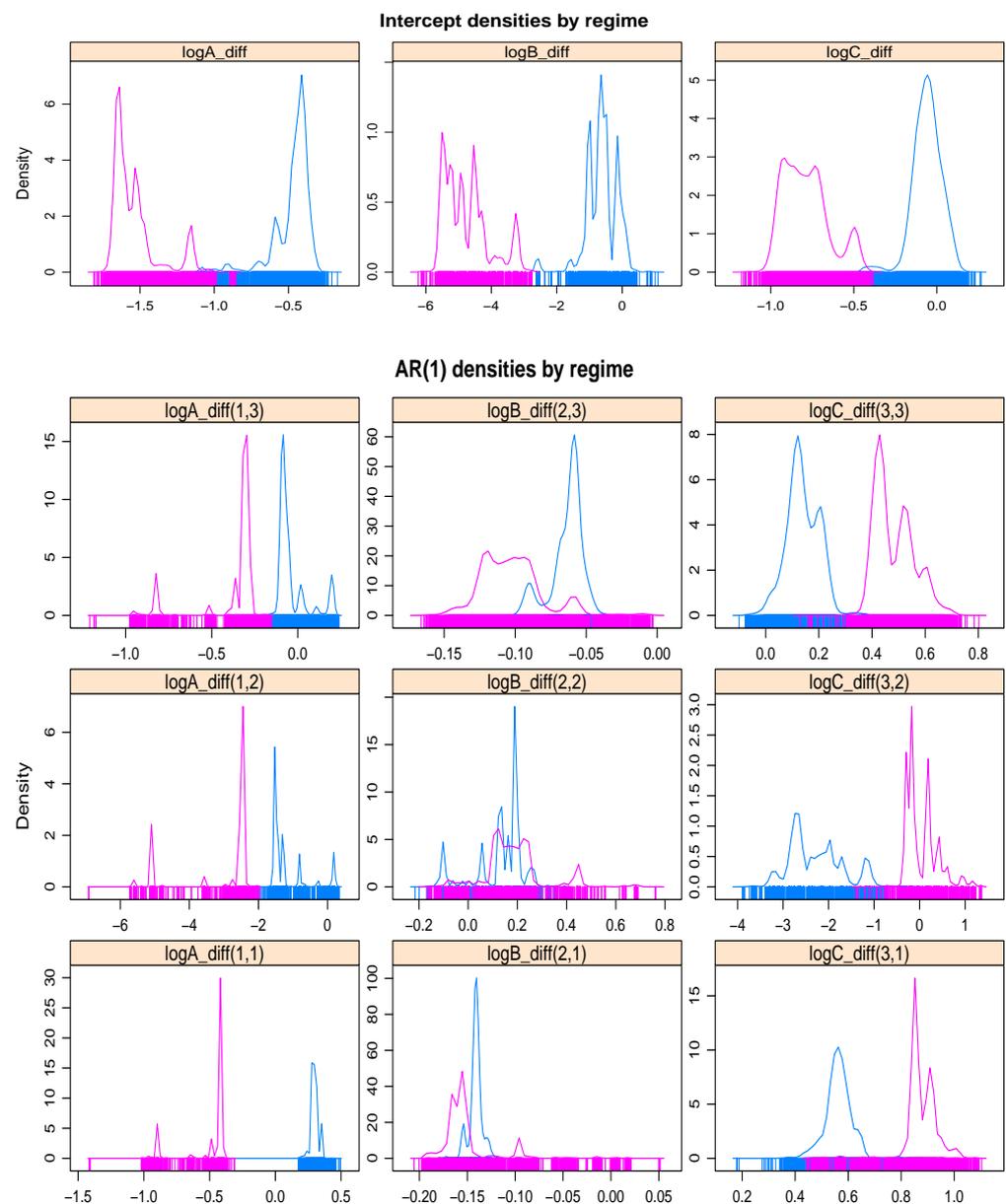


Figure 13. Intercepts and AR(1) densities by regime for Group 1 parameters (transformed) forecasted in the three-year mortality forecast from two-group MSBVAR(1) trained on 1953–2020 French male civilian population data. The pink in the intercept graphs correspond to one regime, and the blue corresponds to the other, and similarly for the AR(1) graphs.

To further compare MSBVAR(1) and BVAR(1), we fit the Group 1 transformed parameters (scaled by 100) with BVAR(1) using the same prior and the same strategy for choosing hyperparameter values described earlier in this section for MSBVAR(1).

Table 10 provides the fitted VAR coefficients and posterior residual covariance matrix from BVAR(1) and MSBVAR(1). The MSBVAR(1) results show that the intercepts and AR coefficients in different regimes are significantly different, which matches the apparent separation between the distribution modes for different regimes in most of the panels in Figure 13.

Table 10. Intercepts, AR coefficients, and error covariance fitted by BVAR(1) and MSBVAR(1) for Group 1 parameters (transformed) in the one-year through three-year mortality forecast by obtained by training on 1953–2020 French male civilian population data.

	BVAR(1)		MSBVAR(1)						
			(The Mean of the 20,000 Draws from Posterior Forecast Density)						
			The 1st Regime			The 2nd Regime			
Intercepts	−2.05862	−4.28405	−0.85628	−0.46341	−0.58897	−0.05221	−1.52664	−4.69968	−0.77721
AR coefficients	−0.31006	−2.00789	−0.25976	0.29793	−1.17564	−0.03485	−0.48458	−2.86573	−0.37377
Posterior residual covariance	0.62313	0.08462	0.14539	0.56322	−2.27056	0.14019	0.87628	0.04168	0.48325
	120.56879	448.4988	76.79942	51.88254	282.05098	35.82726	180.67741	594.69937	113.10181
	448.49879	2552.0473	417.17693	282.05098	2453.57520	318.34130	594.69937	2558.24889	491.59018
	76.79942	417.1769	77.05636	35.82726	318.34130	49.05973	113.10181	491.59018	105.75571

By comparing the posterior residual covariance matrices in Table 10 with the uniform residual covariance from by BVAR(1) for the entire period, MSBVAR(1) can partition the period and granularly assign each part to either a high-volatility regime (the second regime with greater posterior residual covariance) or a low-volatility regime (the first regime with smaller posterior residual covariance). Due to its inability to recognize these regimes, BVAR(1) may overestimate the volatility during a low-volatility regime and underestimate the volatility during a high-volatility regime. Furthermore, MSBVAR(1) can forecast these regimes for future periods; incorporating both regimes in its forecasts can allow for more proper estimation of VAR coefficients and error covariance matrices based on the inclusion of regimes in the mortality forecasts.

Table 11 shows the forecast mean and quantiles for the Group 1 parameters (transformed) from BVAR(1) and MSBVAR(1). By considering two regimes that separate the periods of low and high volatility, MSBVAR(1) can provide quite different mean forecasts and width of the 68% forecast confidence intervals, with a majority of confidence intervals being significantly narrower than those of BVAR(1).

Table 11. Forecast quantiles and means for Group 1 parameters (transformed) from BVAR(1) and MSBVAR(1) in one-year through three-year mortality forecasting after training on 1953–2020 French male civilian population data. For each forecast year, the model that provides a clearly narrower 68% confidence interval is marked in green. As there is no clear difference between the confidence intervals for the two-year forecasting of logC_diff, there is no green mark for this case.

Model	Time Series (Scaled by 100) No. Years Ahead Forecasted	logA_diff			logB_diff			logC_diff		
		1-Year	2-Year	3-Year	1-Year	2-Year	3-Year	1-Year	2-Year	3-Year
		BVAR(1)	16% quantile	−24.40069	−50.68266	−43.10678	72.96510	−128.42168	−82.78513	13.93015
	mean	15.31044	−7.63767	0.23389	107.23746	−39.49984	12.75223	16.97254	−7.49284	1.91958
	84% quantile	54.81178	35.26181	42.70323	140.9184	49.12812	107.46391	19.97902	4.69020	15.67848
MSBVAR(1)	16% quantile	3.761618	−23.17480	−11.93503	45.70565	−114.48584	−48.72999	8.20572	−20.95895	−8.49605
	mean	20.57650	−8.34501	2.44477	111.76375	−47.43908	16.82710	20.92243	−8.66176	3.25535
	84% quantile	37.15495	6.60100	16.32605	176.92057	20.27684	80.27992	33.62026	3.69373	14.73931

4. Conclusions

The approach to mortality modeling in this paper was chosen with two main goals: objectively and automatically capturing and quantifying structural changes in mortality using regime-switching models while retaining the advantages afforded by BVAR. Both of these goals were accomplished with MSBVAR as the forecasting model.

We chose DLGC for the underlying parametric model, rather than, for instance, the Heligman–Pollard (HP) model, because DLGC has better accuracy, especially in recent decades (see Figure 8) and clearer parameter interpretations, and because its parameter time plots can reveal structural changes (Fu et al. 2022). It is worth considering that there are several qualities of the DLGC that make it superior to HP for this application. In future work, we plan to try to develop a more granular understanding of the relative contributions of regime-switching versus the choice of age-specific model on the improvement in fit with

MSBVAR+DLGCab over BVAR+HP. Furthermore, we want to try to understand better which of the good properties of DLGC contributed most to its effectiveness in the present study. More generally, we may try to understand the effects of tuning the age-specific model (for instance, by selecting other variants of DLGC) on forecast quality and the ability to describe structural changes.

The MSBVAR forecasts were better than Lee–Carter, as well as forecasts made with BVAR using the DLGCab model. Our MSBVAR model had an MSE less than half that of the comparable BVAR while having fewer parameters. This is evidence that MSBVAR forecasts can better account for structural changes in historical mortality. This claim is further supported by the structural change identified by MSBVAR around 1990 for later adult ages only.

MSBVAR retains the advantages of BVAR in that it can produce confidence intervals for forecast cross-sectional mortality curves and can provide more realistic and flexible confidence intervals for the forecast parameter values (see Section 3.3.2). However, the addition of regime-switching provides even greater ability. There is a clear separation into different regimes for most parameters when comparing the distribution within each regime, and we found that for most parameters one regime had substantially lower volatility than the other. In addition, we saw unreasonable confidence intervals (i.e., outside the range of possible parameter values) for certain parameters when using BVAR, which was not seen with MSBVAR. Considering the presence of regime-switching, this provides further motivation for choosing MSBVAR over BVAR as the forecasting method.

One concern is that two of the fifteen DLGC parameters, M_2 and $\log T_m$, appeared to undergo permanent structural breaks around 1990 (see Section 3.1). The $\log T_m$ parameter exhibited a stable trend on either side of the break, while M_2 appeared to undergo regime-switching in addition to the break. The parameter M_2 was another exceptional parameter, being weakly correlated with all other parameters, both those that describe old age and those that describe the youngest ages. These factors, while potentially able to reduce the utility of the model, seem to not have had any very detrimental effects. Forecasting accuracy was excellent, and confidence intervals behaved reasonably. We speculate that the apparently successful use of MSBVAR in the presence of structural breaks is because these breaks occurred in only two out of the fifteen parameters, and even M_2 seemed to undergo additional regime switches outside of the break. The remaining parameters all showed excellent behavior.

In future work, we would like to refine our process by using a model that distinguishes regime-switching and structural breaks. In addition, we would like to use break-point testing analysis to automatically determine exactly which coefficients are undergoing regime-switching at a particular moment and to distinguish these from ones that can be modeled by a single regime.

Author Contributions: Conceptualization: W.F.; methodology: W.F.; software: W.F. and S.D.; validation: W.F.; formal analysis: W.F., P.B. and B.R.S.; investigation: W.F.; writing—original draft preparation: W.F. and B.R.S.; writing—review and editing: W.F., P.B. and B.R.S.; supervision: W.F. and P.B.; project administration: W.F. and P.B.; visualization: W.F. and S.D.; funding acquisition: P.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available in a publicly accessible repository that does not issue DOIs Only publicly available datasets were analyzed in this study. The data can be found at: Human Mortality Database, University of California, Berkeley, USA and Max Planck Institute for Demographic Research, Germany, available at www.mortality.org or www.humanmortality.de (data downloaded on 25 August 2022).

Conflicts of Interest: The authors declare no conflict of interest.

Notes

- ¹ Strategy for predetermining hyperparameters in our use of MSBVAR: for Group 1 parameters, we used the hyperparameters values: $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_3 = 1$, $\lambda_4 = 0.1$, $\lambda_5 = 0$, $\mu_5 = 0$, $\mu_6 = 0$ with Sims-Zha normal-flat prior. For Group 2 parameters, among the 46-year training data we fit BVAR on the first 41 years' data and forecast the mortality for the last 5 years with different hyperparameter values. We estimated the posterior and in-sample fit RMSE for the results and picked the one set of hyperparameter values with the lowest RMSE, which was used to set the hyperparameter values in the associated out-of-sample forecast (training on the 46-year data and predicting 1, 2, and 3 years ahead). Here, we set $\lambda_3 = 1$, $\lambda_5 = 0$, $\mu_5 = 0$, and $\mu_6 = 0$, used the Sims-Zha normal-flat prior, and tried several possible values of λ_0 , λ_1 , and λ_4 .
- ² Strategy for predetermining hyperparameters in the BVAR(1) model: we applied the same strategy described in the above footnote for the Group 2 parameters in our model. However, following Njenga and Sherris (2020), we set $\mu_5 = 5$ and $\mu_6 = 5$, which was because the parameter time series were not transformed to be stationary and may have had differences in terms of their trends or stationarity.

References

- Avraam, Demetris, Séverine Arnold, Dyfan Jones, and Bakhtier Vasiev. 2014. Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population. *Experimental Gerontology* 60: 18–30. [\[CrossRef\]](#)
- Bardoutsos, Anastasios, Joop de Beer, and Fanny Janssen. 2018. Projecting delay and compression of mortality. *Genus* 74: 17. [\[CrossRef\]](#)
- Brouhns, Natacha, Michel Denuit, and Jeroen K. Vermunt. 2002. A Poisson Log-Bilinear Regression Approach to the Construction of Projected Life Tables. *Insurance: Mathematics and Economics* 31: 373–93.
- de Beer, Joop, and Fanny Janssen. 2016. A new parametric model to assess delay and compression of mortality. *Population Health Metrics* 14: 1–21. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fu, Wanying, Barry R. Smith, Patrick Brewer, and Sean Droms. 2022. A New Mortality Framework to Identify Trends and Structural Changes in Mortality Improvement and Its Application in Forecasting. *Risks* 10: 161. [\[CrossRef\]](#)
- Gao, Huan, Rogemar Mamon, Xiaoming Liu, and Anton Tenyakov. 2015. Mortality modelling with regime-switching for the valuation of a guaranteed annuity option. *Insurance: Mathematics and Economics* 63: 108–20. [\[CrossRef\]](#)
- Gyls, Rokas, and Jonas Šiaulyš. 2020. Estimation of Uncertainty in Mortality Projections Using State-Space Lee-Carter Model. *Mathematics* 8: 1053. [\[CrossRef\]](#)
- Hainaut, Donatien. 2012. Multidimensional Lee-Carter model with switching mortality processes. *Insurance: Mathematics and Economics* 50: 236–46. [\[CrossRef\]](#)
- Heligman, Larry, and John Pollard. 1980. The age pattern of mortality. *Journal of the Institute of Actuaries* 107: 49–80. [\[CrossRef\]](#)
- Human Mortality Database. n.d. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available online: www.mortality.org (accessed on 25 August 2022).
- Ignatieva, Katja, Andrew Song, and Jonathan Ziveyi. 2016. Pricing and hedging of guaranteed minimum benefits under regime-switching and stochastic mortality. *Insurance: Mathematics and Economics* 70: 286–300.
- Krolzig, Hans-Martin. 2000. *Predicting Markov-Switching Vector Autoregressive Processes*. Nuffield College, University of Oxford, Economics Discussion Papers 2000-W31. Oxford: University of Oxford.
- Lee, Ronald D., and Lawrence R. Carter. 1992. Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* 87: 659–71. [\[CrossRef\]](#)
- Lin, X. Sheldon, and Xiaoming Liu. 2007. Markov Aging Process and Phase-Type Law of Mortality. *North American Actuarial Journal* 11: 92–109. [\[CrossRef\]](#)
- Litterman, Robert B. 1986. Forecasting with Bayesian Vector Autoregressions: Five Years of Experience. *Journal of Business & Economic Statistics* 4: 25–38.
- Lu, Yang, and Dan Zhu. 2023. Modelling Mortality: A bayesian factor-augmented var (favar) approach. *Astin Bulletin* 53: 29–61. [\[CrossRef\]](#)
- McNown, Robert, and Andrei Rogers. 1989. Forecasting mortality: A parameterized time series approach. *Demography* 26: 645–60. [\[CrossRef\]](#)
- Milidonis, Andreas, Yijia Lin, and Samuel H. Cox. 2011. Mortality Regimes and Pricing. *North American Actuarial Journal* 15: 266–89. [\[CrossRef\]](#)
- Njenga, Carolyn Ndigwako, and Michael Sherris. 2020. Modeling mortality with a Bayesian vector autoregression. *Insurance: Mathematics and Economics* 94: 40–57. [\[CrossRef\]](#)
- Robertson, John, and Ellis Tallman. 1999. Vector autoregressions: Forecasting and reality. *Economic Review* 84: 4–18.
- Shen, Yang, and Tak Siu. 2013. Longevity bond pricing under stochastic interest rate and mortality with regime-switching. *Insurance: Mathematics and Economics* 52: 114–23. [\[CrossRef\]](#)
- Sims, Christopher A., Daniel F. Waggoner, and Tao Zha. 2008. Methods for inference in large multiple-equation Markov-switching models. *Journal of Econometrics* 146: 255–74. [\[CrossRef\]](#)
- Sims, Christopher A., and Tao Zha. 1998. Bayesian Methods for Dynamic Multivariate Models. *International Economic Review* 39: 949–68. [\[CrossRef\]](#)

- Thatcher, A. Roger. 1999. The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162: 5–43. [[CrossRef](#)]
- Thompson, Patrick A., William R. Bell, John F. Long, and Robert B. Miller. 1989. Multivariate time series projections of parameterized age-specific fertility rates. *Journal of the American Statistical Association* 84: 689–99. [[CrossRef](#)]
- Zhou, Rui. 2019. Modelling mortality dependence with regime-switching copulas. *ASTIN Bulletin: The Journal of the IAA* 49: 373–407. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.