

RESEARCH PAPER

The impact of long memory in mortality differentials on index-based longevity hedges

Kenneth Q. Zhou¹ and Johnny Siu-Hang Li^{2*}

¹School of Mathematical and Statistical Sciences, Arizona State University, USA and ²Department of Statistics and Actuarial Science, University of Waterloo, Canada

*Corresponding author: E-mail: shli@uwaterloo.ca

(Received 6 April 2023; revised 6 April 2023; accepted 6 April 2023)

Abstract

In multi-population mortality modeling, autoregressive moving average (ARMA) processes are typically used to model the evolution of mortality differentials between different populations over time. While such processes capture only short-term serial dependence, it is found in our empirical work that mortality differentials often exhibit statistically significant long-term serial dependence, suggesting the necessity for using long memory processes instead. In this paper, we model mortality differentials between different populations with long memory processes, while preserving coherence in the resulting mortality forecasts. Our results indicate that if the dynamics of mortality differentials are modeled by long memory processes, mean reversion would be much slower, and forecast uncertainty over the long run would be higher. These results imply that the true level of population basis risk in index-based longevity hedges may be larger than what we would expect when ARMA processes are assumed. We also study how index-based longevity hedges should be calibrated if mortality differentials follow long memory processes. It is found that delta hedges are more robust than variance-minimizing hedges, in the sense that the former remains effective even if the true processes for mortality differentials are long memory ones.

Keywords: ARFIMA processes; longevity Greeks; population basis risk; the Li–Lee model

1. Introduction

To mitigate longevity risk exposures, pension plans and annuity providers may deploy index-based longevity hedges. An index-based longevity hedge is constructed using one or more instruments whose payoffs are linked to a mortality index which tracks the mortality experience of a certain reference population, typically a national population. Researchers have studied index-based longevity hedges from different angles, ranging from the development of effective hedging strategies to the quantification of residual risks that still remains when a properly calibrated index-based longevity hedge is in place [Dahl *et al.* (2008); Cairns (2011); Coughlan *et al.* (2011); Li and Hardy (2011); Cairns *et al.* (2014); Zhou and Li (2017); Li *et al.* (2021)].

© Université catholique de Louvain 2023. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the residual risks is population basis risk, which arises due to the difference in mortality experience between the hedger's population of individuals and the population to which the hedging instrument is linked. The quantification of population basis risk requires a multi-population mortality model, which models the mortality dynamics of the populations involved jointly. Examples of such models include those proposed by Cairns *et al.* (2011), Dowd *et al.* (2011), Zhou *et al.* (2014), Kleinow (2015) and Enchev *et al.* (2017). Reviews of such models are provided by Li *et al.* (2015), and Villegas *et al.* (2017).

In multi-population mortality models, the evolution of mortality differentials between populations is often captured by autoregressive moving average (ARMA) process. As ARMA processes are stationary, the long-term expected mortality differential between any two populations being modeled is finite. In effect, the divergence between the expected mortality trajectories of two related populations is bounded. This property is commonly referred to as "coherence," and is regarded as desirable. However, since ARMA processes are short memory processes, they do not capture any long-term serial dependence in mortality differentials. This limitation may potentially affect the assessment of population basis risk and hedge effectiveness.

Recent studies have shown that long-term serial dependence exists in mortality dynamics. Yan *et al.* (2021) showed empirically the existence of long memory in age-specific mortality rates from 16 countries, and proposed a long memory mortality model that is based on the generalized linear modeling approach. Yan *et al.* (2020) extended the work of Yan *et al.* (2021) to multivariate mortality modeling with cohort effects. Using a continuous-time setting, Wang *et al.* (2021) proposed the Volterra mortality model with long-range dependence. Based on the Volterra model, Wang and Wong (2021) developed a time-consistent mean-variance longevity hedge. Other studies related to long memory in mortality dynamics include those of Gil-Alana *et al.* (2017), Delgado-Vences and Ornelas (2019) and Yaya *et al.* (2019).

The aforementioned studies studied long-term serial dependence in mortality dynamics but not mortality differentials that determine population basis risk. To fill this gap, in this paper we attempt to verify the existence of long memory in mortality differentials, and investigate the impact of this property on index-based longevity hedges. To this end, we consider modeling mortality differentials between related populations with autoregressive fractionally integrated moving average (ARFIMA) processes [Granger and Joyeux (1980)], which may be seen as autoregressive integrated moving average (ARIMA) processes with a fractional differentiation. The long-term mean implied by an ARFIMA process is finite, so that the coherence property can be preserved when ARMA is replaced with ARFIMA in the modeling of mortality differentials. More importantly, an ARFIMA process can simultaneously capture both long- and short-term serial dependence in mortality differentials via its fractional differentiation and ARMA components, respectively. Although Hyndman *et al.* (2013) previously used ARFIMA processes for mortality modeling, their work was based on a rather different modeling approach (the product ratio method) and took no consideration of any actuarial application.

The contributions of this paper are twofold. First, using mortality data from eight national populations, we demonstrate that mortality differentials between populations often exhibit statistically significant long-term serial dependence. To capture this feature and replicate it in mortality forecasts, we use an ARFIMA process to model the evolution of each population-specific period effect in the Li-Lee (LL) model

[Li and Lee (2005)]. Our empirical analysis reveals that compared to an ARMA process that takes no account of long-term serial dependence, an ARFIMA process implies a slower rate of mean reversion in mortality differentials and a higher level of forecast uncertainty over the long run. We also show that the proposed modeling approach preserves the coherence property.

Second, we study how long memory in mortality differentials may affect index-based longevity hedges. Specifically, we first adapt existing delta-neutral and variance-minimizing hedging strategies to fit our modeling framework, and then present two case studies that focus on hedge effectiveness and robustness of hedging strategies, respectively. The first case study reveals that overlooking long memory would lead to an underestimation of population basis risk and consequently an overestimation of hedge effectiveness. The second case study points to the conclusion that the hedge ratio (and hedge effectiveness) of a variance-minimizing hedge depends heavily on whether long memory processes are utilized to model mortality differentials in the course of calibration, but the opposite is true for a delta-neutral hedge.

The rest of this paper is organized as follows. Section 2 examines the statistical significance of long memory in mortality differentials, and presents the ARFIMA modeling work for the dynamics of mortality differentials (population-specific period effects in the LL model) over time. Section 3 studies how existing hedging strategies should be modified when the proposed modeling approach is used. Section 4 features two case studies that highlight the implications of long memory in mortality differentials on index-based longevity hedges. Lastly, section 5 concludes the paper.

2. Modeling long memory in mortality differentials

2.1. Data

We consider mortality data from eight female populations in Europe (see Table 1), provided by the Human Mortality Database. For consistency reasons, the same sample period (1900–2018) and age range (40–89) are used for all populations under consideration. The age range includes typical ages of pension plan members (active and retired), but excludes extreme ages for which data are often extrapolated and smoothed. We set the beginning point of the sample period to 1900, as a longer time-series facilitates the study of long memory.

Table 1. A summary of the mortality data used in this paper

Population	Abbreviation	Sample period	Age range	Gender
England and Wales	EW	1900–2018	40–89	Female
France	FR	1900–2018	40–89	Female
Switzerland	CH	1900–2018	40–89	Female
Denmark	DK	1900–2018	40–89	Female
Finland	FI	1900–2018	40–89	Female
Italy	IT	1900–2018	40–89	Female
Netherlands	NL	1900–2018	40–89	Female
Sweden	SE	1900–2018	40–89	Female

2.2. Mortality model

Our modeling work is based on the well-known Li–Lee (LL) model [Li and Lee (2005)]. Let $m_{x,t}^{(i)}$ be the central rate of death at age x in year t for population i , where $i \in \mathbb{P}$ with \mathbb{P} being the set of populations under consideration. The LL model is specified as

$$\ln m_{x,t}^{(i)} = a_x^{(i)} + B_x K_t + b_x^{(i)} k_t^{(i)} + e_{x,t}^{(i)}, \tag{1}$$

where

- $a_x^{(i)}$ is an age-specific parameter indicating the i -th population’s average level of mortality at age x ,
- K_t is a time-varying index that is shared by all populations in \mathbb{P} ,
- B_x is an age-specific parameter indicating the sensitivity of $\ln(m_{x,t}^{(i)})$ to K_t ,
- $k_t^{(i)}$ is a time-varying index that is specific to the i -th population,
- $b_x^{(i)}$ is an age-specific parameter indicating the sensitivity of $\ln(m_{x,t}^{(i)})$ to $k_t^{(i)}$, and
- $e_{x,t}^{(i)}$ is the error term that captures all remaining variations.

It is well-known that the LL model is subject to an identifiability problem. To stipulate parameter uniqueness, the following constraints are used:

$$\begin{aligned} \sum_x B_x &= 1, & \sum_t K_t &= 0, \\ \sum_x b_x^{(i)} &= 1, & i \in \mathbb{P} \text{ and } \sum_t k_t^{(i)} &= 0, \quad i \in \mathbb{P}, \end{aligned}$$

where the summations are taken over the whole sample period or age range. Following Li and Lee (2005), we estimate the parameters in Equation (1) using a singular value decomposition.

In applications of the LL model, $\{K_t\}$ is typically assumed to follow a random walk with drift:

$$K_t = \mu + K_{t-1} + \epsilon_t, \tag{2}$$

where μ is the drift term, and ϵ_t is the time- t random innovation that is normally distributed with a zero mean and a constant standard deviation σ .

On the other hand, $\{k_t^{(i)}\}$ is often assumed to follow an ARMA(P,Q) process:

$$\Phi_p^{(i)}(B)k_t^{(i)} = \Theta_Q^{(i)}(B)\epsilon_t^{(i)}, \tag{3}$$

where $\epsilon_t^{(i)}$ is the time- t random innovation that is normally distributed with a zero mean and a constant standard deviation $\sigma^{(i)}$, $\Phi_p^{(i)}(B) = 1 - \phi_1^{(i)}B - \phi_2^{(i)}B^2 - \dots - \phi_p^{(i)}B^p$, and $\Theta_Q^{(i)}(B) = 1 - \theta_1^{(i)}B - \theta_2^{(i)}B^2 - \dots - \theta_Q^{(i)}B^Q$, with B being the backshift operator (i.e., $Bk_t^{(i)} = k_{t-1}^{(i)}$) and the roots of both $\Phi_p^{(i)}(B)$ and $\Theta_Q^{(i)}(B)$ lying outside the unit circle. Given how the ARMA model is specified, the long-term mean of $k_t^{(i)}$ is finite and equals zero.¹ As a result, the long-term expected mortality differential (in log scale)

¹As the identifiability constraint $\sum_t k_t^{(i)} = 0$ is imposed, the mean of $k_t^{(i)}$ over the sample period is zero; hence, it is appropriate to model $\{k_t^{(i)}\}$ using a process with a zero unconditional mean. The long-term

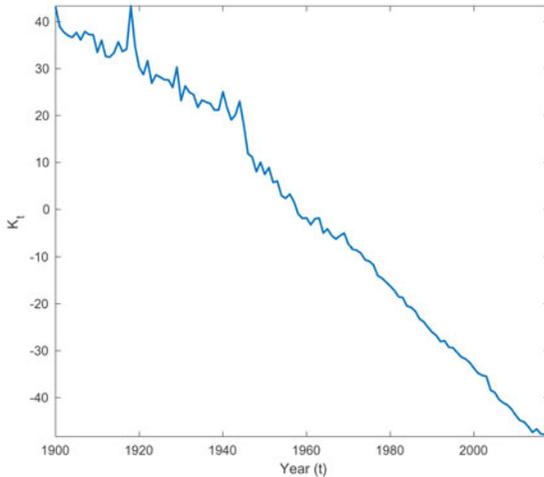


Figure 1. Estimates of K_t within the sample period of 1900–2018.

between populations i and j is

$$\ln m_{x,t}^{(i)} - \ln m_{x,t}^{(j)} = a_x^{(i)} - a_x^{(j)}, \tag{4}$$

a constant that is free of t , so that the coherence property is achieved.

In the rest of this subsection, we investigate whether long memory exists in $\{K_t\}$ and $\{k_t^{(i)}\}$, and suggests suitable processes for modeling $\{K_t\}$ and $\{k_t^{(i)}\}$ accordingly.

2.2.1. Dynamics of K_t

Figure 1 shows the estimates of K_t , the time-varying index shared by the eight populations under consideration. Given the steady trend in the estimates, we apply first differencing and examine whether long memory exists in $\{K_t - K_{t-1}\}$ using the modified R/S test [Lo (1991)], of which the null hypothesis is that the underlying process is a short memory one.

Reported in Table 2, the values of the R/S test statistic indicate that the null hypothesis cannot be rejected for lags 0–5, suggesting that long memory is not significant in $\{K_t - K_{t-1}\}$.²

Given the results of the modified R/S test, we use a random walk with drift to model $\{K_t\}$ throughout the rest of this paper. While more sophisticated processes, such as those that feature conditional heteroskedasticity [Zhou and Li (2020)], may provide a better fit to $\{K_t\}$, we choose to use a simple random walk with drift so that our discussions can focus on the issue of short and long memory. Furthermore, as K_t is shared by all populations under consideration, the process for K_t has no direct relevance to population basis risk.

mortality differential between populations i and j , where $i \neq j$, is captured by the difference between parameters $a_x^{(i)}$ and $a_x^{(j)}$, as indicated in Equation (4).

²According to Lo (1991), when the sample size is small (100–250 observations), the power of the modified R/S test will be low if the number of lags included in calculating the modified R/S statistics is large. Given that our sample size is 118, lags higher than 5 are excluded.

Table 2. Values of the modified R/S test statistic for $\{K_t - K_{t-1}\}$ at lags 0, 1, ..., 5

Number of lags	0	1	2	3	4	5
R/S statistic	0.7719	0.9475	1.0783	1.2117	1.1653	1.2335

At 2.5%, 5%, and 10% significance levels, the critical values of the modified R/S test are 1.862, 1.747, and 1.620, respectively, according to Lo (1991).

2.2.2. Dynamics of $k_t^{(i)}$

Figure 2 shows the estimates of $k_t^{(i)}$ for the eight populations under consideration. Compared to $\{K_t\}$, the trends in $\{k_t^{(i)}\}$ are not as apparent. To better understand the dynamics of $\{k_t^{(i)}\}$, we consider the sample autocorrelation functions (ACF) of $\{k_t^{(i)}\}$. The lag- k sample ACF of $\{k_t^{(i)}\}$ is defined as the sample correlation between $k_t^{(i)}$ and $k_{t-k}^{(i)}$ over the sample period.

The nature of a time-series can be told from its sample ACF plot (the plot of sample ACF against lag). In particular, the ACF of a long memory time series converges more slowly than that of a short memory one (which decays exponentially), but is not as extreme as that of an integrated (difference stationary) process which does not converge. To illustrate, in Figure 3 we display the sample ACF plots of time-series generated from three processes: (a) an AR(1) process with an autoregressive parameter of 0.7 and a volatility parameter of 1, (b) an ARFIMA(1, d ,0) with $d=0.4$ and the same autoregressive and volatility parameters as those of (a), and (c) an ARIMA(0,1,0) process with the same volatility parameter as the other two processes. It can be observed that the ACF of the ARFIMA-generated series clearly has a lower convergence rate than that of the AR-generated series, while the ACF of the ARIMA (0,1,0) process does not converge.

Figure 4 shows the sample ACF plots of the estimated series of $k_t^{(i)}$ for the eight populations under consideration. All of the eight sample ACF plots behave like that of a typical ARFIMA process that features long memory.

We further use the modified R/S test to confirm the existence of long memory in the eight estimated series of $k_t^{(i)}$. As shown in Table 3, for all lags up to 5, the values of modified R/S test statistic are strictly greater than the critical value at 5% significance

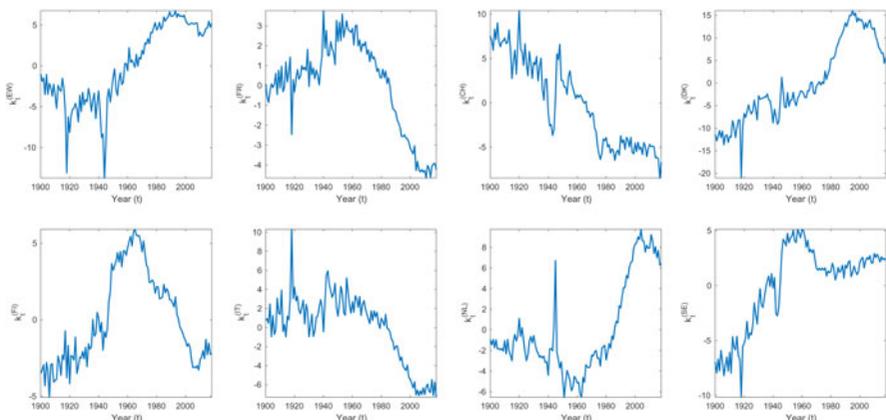


Figure 2. Estimates $k_t^{(i)}$ for all $i \in \mathbb{P}$ within the sample period of 1900–2018.

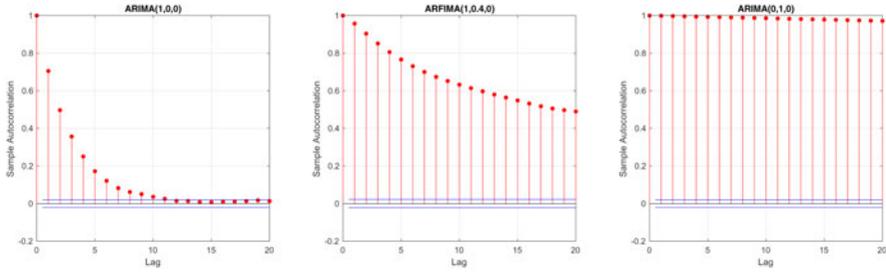


Figure 3. Sample ACF plots of time-series generated from an AR(1) process (left panel), an ARFIMA(1, 0.4, 0) (middle panel), and an ARIMA(0, 1, 0) process (right panel); the autoregressive parameter for the first two processes is 0.7, and the volatility parameter for all three processes is 1.

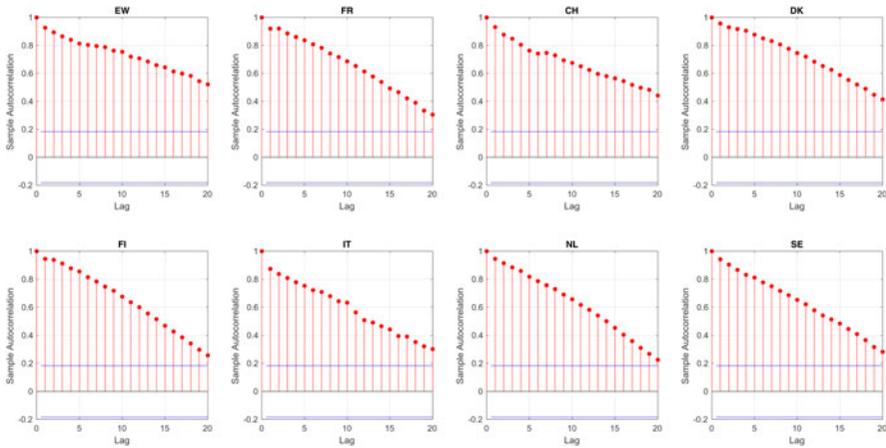


Figure 4. Sample ACF plots of the estimated series of $k_t^{(i)}$ for the eight populations under consideration.

level, indicating that long memory in the eight estimated series of $k_t^{(i)}$ is statistically significant.

2.3. The ARFIMA process

To capture the long memory in $k_t^{(i)}$, we consider an ARFIMA process for $k_t^{(i)}$:

$$\Phi_p^{(i)}(B)(1 - B)^{2d}k_t^{(i)} = \Theta_Q^{(i)}(B)\epsilon_t^{(i)}, \tag{5}$$

where d is the fractional difference parameter, $\epsilon_t^{(i)}$ is the time- t innovation which follows a normal distribution with a zero mean and a standard deviation of $\sigma^{(i)}$, and $\Phi_p^{(i)}(B)$ and $\Theta_Q^{(i)}(B)$ are the autoregressive and moving average operators as defined in the previous subsection, respectively. As the identifiability constraint $\sum_t k_t^{(i)} = 0$ for all $i \in \mathbb{P}$ is used, we consider an ARFIMA specification that has a zero unconditional mean.

The ARFIMA process captures the long-term serial dependence in $\{k_t^{(i)}\}$ through the fractional difference parameter d , and the short-term serial dependence in $\{k_t^{(i)}\}$ via the autoregressive and moving-average operators. The fractional difference parameter d can

Table 3. Values of the modified R/S test statistic for $\{k_t^{(i)}\}$ at lags 0, 1, ..., 5

Population	Number of lags					
	0	1	2	3	4	5
EW	4.738	3.413	2.816	2.458	2.213	2.033
FR	4.236	3.058	2.514	2.190	1.969	1.807
CH	4.637	3.337	2.758	2.411	2.176	2.003
DK	4.583	3.277	2.694	2.344	2.105	1.929
FI	4.775	3.424	2.811	2.446	2.198	2.016
IT	4.184	3.057	2.536	2.220	2.004	1.844
NL	4.293	3.078	2.534	2.209	1.988	1.826
SE	4.462	3.202	2.639	2.304	2.076	1.908

At 2.5%, 5%, and 10% significance levels, the critical values of the modified R/S test are 1.862, 1.747, and 1.620, respectively, according to Lo (1991).

take any real value between -0.5 and 0.5 inclusive. If $d = 0$, then the ARFIMA process reduces to an ARMA process with short-term memory only. If $d = 0.5$, then Equation (5) becomes an ARIMA process with first-order differencing. For $d \in (0, 0.5)$, the ARFIMA process is said to have long memory or persistence, which implies that a high value of $k_t^{(i)}$ is likely to be followed by high values $k_s^{(i)}$ for $s > t$ over a prolonged period of time. Finally, if $d \in (-0.5, 0)$, then the process is said to have anti-persistence, which means that the series will likely to switch between high and low values in adjacent time points for a prolonged period of time.

2.4. Estimation

We estimate AR(1), ARMA(P, Q), and ARFIMA(P, d, Q) processes to the eight estimated series of $k_t^{(i)}$, using R package “forecast.” We consider AR(1), as it is frequently used to model $\{k_t^{(i)}\}$ in the LL model including the original work of Li and Lee (2005). When fitting an ARMA(P, Q) process, the optimal values of P and Q are identified by the “auto.arima” function; and when fitting an ARFIMA(p, d, q) process, the optimal values of P , d , and Q are found using the “arfima” function.

In what follows, we report the estimation results for English and Welsh female population (EW). The estimation results for the other seven populations under consideration have similar properties, and are therefore not shown for the sake of space. Table 4 reports the estimates of the parameters in the processes for $\{k_t^{(EW)}\}$. For both ARMA and ARFIMA processes, the optimal values of P and Q are 1. The estimate of d in the ARFIMA(1, d , 1) process is 0.3787, indicating that there is strong long-term serial dependence in $\{k_t^{(EW)}\}$. All of the parameters are significant, as indicated by their p-values.

Figure 5 displays the sample ACF plots of the residuals from the three fitted processes for $\{k_t^{(EW)}\}$. For the AR process, the sample ACF plot of the residuals has significant spikes at lag 1 and 10, indicating that some short-term serial dependence is not adequately captured by the process. Both the ARMA and ARFIMA processes

Table 4. Estimates of the parameters in the AR, ARMA, and ARFIMA processes for $\{k_t^{(EW)}\}$

Process	Parameter	Estimate	Standard error	p-value
AR(1)	$\phi_1^{(EW)}$	0.9287	0.0321	$<2.2 \times 10^{-16}$
	$\sigma_1^{(EW)}$	1.7564	N/A	N/A
ARMA(1,1)	$\phi_1^{(EW)}$	0.9719	0.0205	$<2.2 \times 10^{-16}$
	$\theta_1^{(EW)}$	-0.3617	0.1029	4.399×10^{-4}
	$\sigma_1^{(EW)}$	1.6822	N/A	N/A
ARFIMA(1,d,1)	d	0.3787	0.0056	$<2.2 \times 10^{-16}$
	$\phi_1^{(EW)}$	0.9682	0.0736	$<2.2 \times 10^{-16}$
	$\theta_1^{(EW)}$	0.7916	0.0305	$<2.2 \times 10^{-16}$
	$\sigma_1^{(EW)}$	1.6716	N/A	N/A

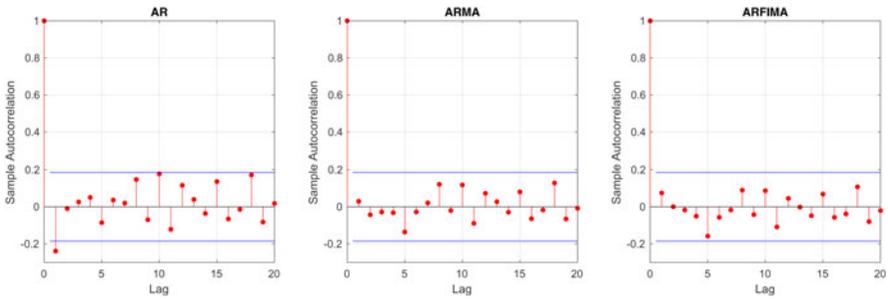


Figure 5. Sample ACF plots of the residuals from the fitted processes for $\{k_t^{(EW)}\}$: AR(1) (left panel), ARMA(1,1) (middle panel), and ARFIMA(1,d,1) process (right panel).

seem to have provided adequate provision for short-term serial dependence, but the former takes no account of long memory which is found to be statistically significant in section 2.2.2.

2.5. Forecasting

We now turn to forecasting. Figure 6 shows the mean forecasts and predictive intervals of $k_t^{(EW)}$ produced by the three estimated processes. The trajectories of the mean forecasts underscore an important property of ARFIMA processes: the rate of convergence implied by an ARFIMA process is smaller than the corresponding AR and ARMA processes. Another important property of ARFIMA processes can be inferred from the width of the predictive intervals: an ARFIMA process implies a higher level of forecast uncertainty compared to the corresponding AR and ARMA processes.

To obtain deeper insights about the rate of convergence, we consider additionally the female population of Italy (IT). For the reader’s information, the estimated AR, ARMA, and ARFIMA processes for $\{k_t^{(IT)}\}$ are summarized in Table 5.

Figure 7 shows the mean forecasts of $\ln m_{x,t}^{(EW)}$ and $\ln m_{x,t}^{(IT)}$ for $x = 40, 50, 60$ when AR, ARMA, and ARFIMA processes are used to model $\{k_t^{(EW)}\}$ and $\{k_t^{(IT)}\}$. It can be

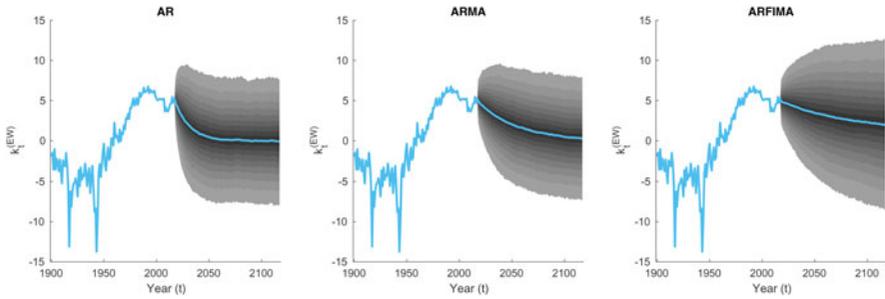


Figure 6. Mean forecasts and predictive intervals of $k_t^{(EW)}$ generated from the fitted AR process (left panel), ARMA process (middle panel), and ARFIMA process (right panel). Each fan chart shows the 10% predictive interval with the heaviest shading, surrounded by the 20%, 30%, ..., 90% predictive intervals with progressively lighter shadings.

observed that all three types of processes (including ARFIMA) produce coherent mortality forecasts; that is, for a given age x , the mean forecasts of $\ln m_{x,t}^{(EW)}$ and $\ln m_{x,t}^{(IT)}$ do not diverge over the long run.

Noting that all of the three processes specified for $k_t^{(i)}$ have a zero unconditional mean, it can be easily deduced that in the long-run equilibrium the difference between the log central death rates for EW and IT populations at any given age x is always $a_x^{(EW)} - a_x^{(IT)}$ regardless of which of the three processes are used. Figure 7 therefore reveals that three types of processes lead to very different rates of convergence to the long-term equilibrium. In particular, the rate of convergence to the long-term equilibrium is the fastest (slowest) when AR (ARFIMA) processes are used to model $\{k_t^{(EW)}\}$ and $\{k_t^{(IT)}\}$.

3. Hedging strategies

Drawing on the previous works of Zhou and Li (2020) and Zhou and Li (2021), we now derive hedging strategies that fit the modeling framework described in earlier sections. We consider both delta-neutral and variance-minimizing index-based longevity hedges.

Table 5. Estimates of the parameters in the AR, ARMA, and ARIMA processes for $\{k_t^{(IT)}\}$

Process	Parameter	Estimate	Standard error	p-value
AR(1)	$\phi_1^{(IT)}$	0.8975	0.0409	$<2.2 \times 10^{-16}$
	$\sigma^{(IT)}$	1.6435	N/A	N/A
ARMA(1,1)	$\phi_1^{(IT)}$	0.9850	0.0156	$<2.2 \times 10^{-16}$
	$\theta_1^{(IT)}$	-0.5414	0.0867	4.196×10^{-10}
	$\sigma^{(IT)}$	1.4950	N/A	N/A
ARFIMA(1,d,1)	d	0.2145	0.0041	$<2.2 \times 10^{-16}$
	$\phi_1^{(IT)}$	0.9813	0.0712	$<2.2 \times 10^{-16}$
	$\theta^{(IT)}$	0.7440	0.0230	$<2.2 \times 10^{-16}$
	$\sigma^{(IT)}$	1.4848	N/A	N/A

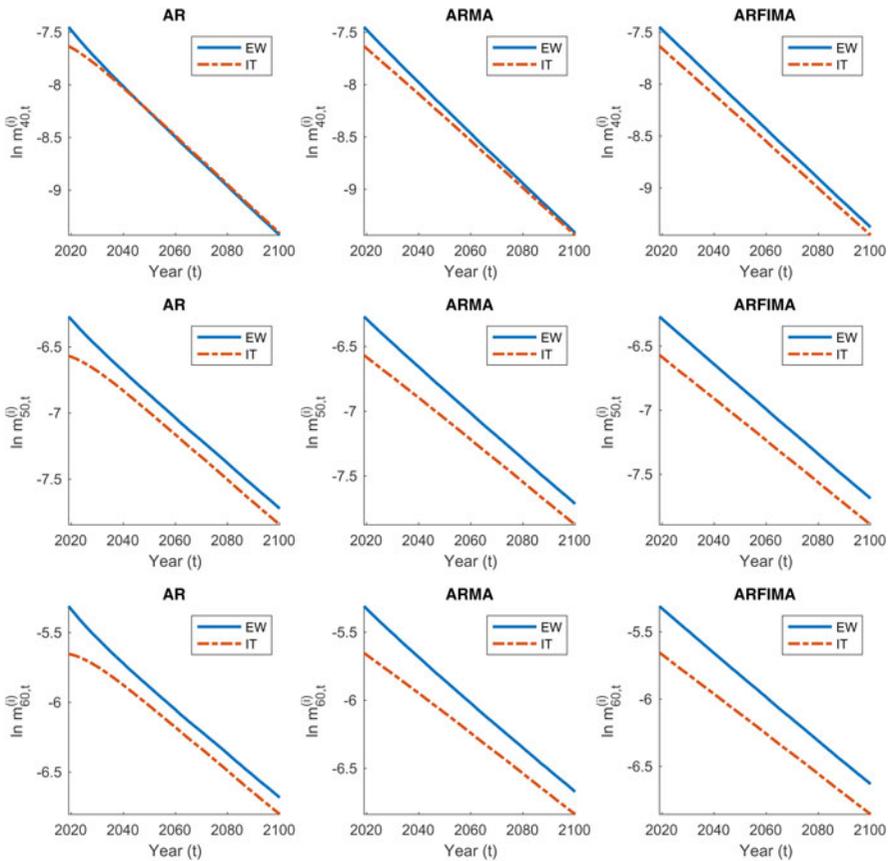


Figure 7. Mean forecasts of $\ln m_{x,t}^{(EW)}$ and $\ln m_{x,t}^{(IT)}$ at $x=40$ (top row), $x=50$ (middle row), and $x=60$ (bottom row) when $\{k_t^{(EW)}\}$ and $\{k_t^{(IT)}\}$ are modeled by AR processes (left column), ARMA processes (middle column), and ARFIMA processes (right column).

3.1. Set-up

Let $S_{x,t}^{(i)}(T)$ be the *ex post* probability that an individual from population i who has survived to age x at time t would have survived to time $t + T$ for $T = 1, 2, \dots$. Under the LL model, $S_{x,t}^{(i)}(T)$ is given by

$$S_{x,t}^{(i)}(T) = \exp\left(-\sum_{s=1}^T \exp\left(a_{x+s-1}^{(i)} + B_{x+s-1}K_{t+s} + b_{x+s-1}^{(i)}k_{t+s}^{(i)}\right)\right). \tag{6}$$

As a shorthand, we write

$$S_{x,t}^{(i)}(T) = e^{-W_{x,t}^{(i)}(T)}, \tag{7}$$

where

$$W_{x,t}^{(i)}(T) = \sum_{s=1}^T \exp(Y_{x,t}^{(i)}(s)) \tag{8}$$

and

$$Y_{x,t}^{(i)}(s) = a_{x+s-1}^{(i)} + B_{x+s-1}K_{t+s} + b_{x+s-1}^{(i)}K_{t+s}^{(i)}. \tag{9}$$

Let \mathcal{F}_t be the information about the evolution of mortality of all populations under consideration up to and including time t . Then, the survival probability for an individual from population i aged x at time t to survive to time $t + T$ given \mathcal{F}_t can be expressed as

$$E\left[e^{-W_{x,t}^{(i)}(T)} \mid \mathcal{F}_t\right].$$

Under the LL model, dependence across populations is driven exclusively by $\{K_t\}$. To stress the role of $\{K_t\}$, we define the following:

$$p_{x,t}^{(i)}(T, K_t) := E\left[e^{-W_{x,t}^{(i)}(T)} \mid \mathcal{F}_t\right].$$

Note that the expression above depends on K_t but not K_{t-1}, K_{t-1}, \dots due to the Markovian property of the assumed process for $\{K_t\}$.

To construct a delta-neutral longevity hedge, we need the longevity delta for $p_{x,t}^{(i)}(T, K_t)$, which is defined as the first-order partial derivatives of $p_{x,t}^{(i)}(T, K_t)$ with respect to K_t :

$$\begin{aligned} \Delta_{x,t}^{(i)}(T) &:= \frac{\partial}{\partial K_t} p_{x,t}^{(i)}(T, K_t) \\ &= - \sum_{s=1}^T B_{x+s-1} E\left[\exp(Y_{x,t}^{(i)}(s) - W_{x,t}^{(i)}(T)) \mid \mathcal{F}_t\right]. \end{aligned}$$

We assume that the hedger is an annuity provider and the hedging instrument is an S-forward. In the next two subsections, we derive the time- t values and longevity deltas of a life annuity and S-forward.

3.1.1. Life annuities

Let us consider a τ -year deferred T -year temporary life annuity issued to an individual from population i who is aged x at time t . Assuming a constant interest rate of r , the sum of all discounted cash flows from this life annuity is given by

$$\mathcal{L}^{(i)} = \sum_{s=1}^T (1+r)^{-(\tau+s)} S_{x,t}^{(i)}(\tau+s).$$

As a linear combination of $S_{x,t}^{(i)}(\tau+s)$ for $s = 1, \dots, T$, the time- t value of $\mathcal{L}^{(i)}$ given \mathcal{F}_t can be written as

$$L^{(i)}(K_t) := E[\mathcal{L}^{(i)} \mid \mathcal{F}_t] = \sum_{s=1}^T (1+r)^{-(\tau+s)} p_{x,t}^{(i)}(\tau+s, K_t).$$

The longevity delta of this life annuity is defined as the first-order partial derivative of $L^{(i)}(K_t)$ with respect to K_t ; that is,

$$\Delta_L^{(i)} := \frac{\partial L^{(i)}(K_t)}{\partial K_t} = \sum_{s=1}^T (1+r)^{-(\tau+s)} \Delta_{x,t}^{(i)}(\tau+s).$$

3.1.2. S-forwards

A S-forward is a zero-coupon swap with a fixed leg proportional to a fixed forward rate S^f that is determined when the contract is written at time t , and a floating leg proportional to a random survival rate. For a S-forward with a reference population i , a time-to-maturity T and a reference age x , the floating leg is proportional to $S_{x,t}^{(i)}(T)$, as defined in Equation (6).

The value of $S_{x,t}^{(i)}(T)$ is completely realized at time $t + T$ when the S-forward matures. If the realized value of $S_{x,t}^{(i)}(T)$ turns out to be higher than expected (i.e., realized mortality is lighter than expected), then the fixed-rate payer of the S-forward receives a positive net payment from the floating-rate payer. To hedge their longevity exposures, pension and annuity providers may participate in a S-forward as a fixed-rate payer, so that in case mortality turns out to lighter than expected, the net payment received from the floating-rate payer can be used to offset their correspondingly larger liabilities.

Assuming a constant interest rate of r , the discounted payoff of the above-described S-forward from the fixed-rate payer’s perspective is given by

$$\mathcal{H}^{(i)} = (1+r)^{-T} (S_{x,t}^{(i)}(T) - S^f)$$

per \$1 notional. Given \mathcal{F}_t , the time- t value of this S-forward can be expressed as

$$H^{(i)}(K_t) := E[\mathcal{H}^{(i)} | \mathcal{F}_t] = (1+r)^{-T} (p_{x,t}^{(i)}(T, K_t) - S^f).$$

The longevity delta is defined as the first-order partial derivative of $H^{(i)}(K_t)$ with respect to K_t ; that is,

$$\Delta_H^{(i)} := \frac{\partial H^{(i)}(K_t)}{\partial K_t} = (1+r)^{-T} \Delta_{x,t}^{(i)}(T).$$

3.2. Calibration and evaluation

Suppose that the liability being hedged is the life annuity described in section 3.1.1 and the hedging instrument is the S-forward described in section 3.1.2. Then, the unhedged position is simply $\mathcal{L}^{(i)}$, while the hedged position can be expressed as $\mathcal{L}^{(i)} - u\mathcal{H}^{(j)}$, where u is the hedge ratio (the notional amount of the S-forward purchased). When $i \neq j$, the life annuity and S-forward are associated with different populations, and thus population basis risk exists. We consider two methods to choose u .

For a delta-neutral hedge established at time t , the value of u is determined such that the longevity delta of the hedged position is zero; that is,

$$\Delta_L^{(i)} - u\Delta_H^{(j)} = 0,$$

the solution to which is given by

$$u_{\Delta}^{(i,j)} = \frac{\Delta_L^{(i)}}{\Delta_H^{(j)}}. \quad (10)$$

For a variance-minimizing hedge established at time t , the value of u is chosen such that the following is minimized:

$$\text{Var}(\mathcal{L}^{(i)} - u\mathcal{H}^{(j)}|\mathcal{F}_t).$$

It can be shown that the solution to the above minimization problem is

$$u_V^{(i,j)} = \frac{\text{Cov}(\mathcal{L}^{(i)}, \mathcal{H}^{(j)}|\mathcal{F}_t)}{\text{Var}(\mathcal{H}^{(j)}|\mathcal{F}_t)}. \quad (11)$$

Finally, to quantify hedge effectiveness, we consider the reduction in variance between the hedged and unhedged positions:

$$\text{HE} = 1 - \frac{\text{Var}(\mathcal{L}^{(i)} - u\mathcal{H}^{(j)}|\mathcal{F}_{t_0})}{\text{Var}(\mathcal{L}^{(i)}|\mathcal{F}_{t_0})}, \quad (12)$$

where u represents the hedge ratio, which is calculated using Equation (10) when a delta-neutral hedge is used and Equation (11) when a variance-minimizing hedge is used.

4. Numerical illustrations

In this section, we present two case studies to demonstrate the impact of long memory in mortality differentials on index-based longevity hedges. Both case studies are based on the data and model discussed in section 2.

4.1. Case study I

The following assumptions are used for case study I:

- The current time is $t_0 = 2018$ (i.e., the end of the last year of the sample period).
- The liability being hedged is a 25-year deferred 25-year temporary life annuity issued to an individual from population i who is aged 40 at time $t_0 = 2018$. The annuity pays \$1 at the end of each year, starting from age 65. Payment ceases when the annuitant dies or reaches age 90, whichever is the earliest. The annuitant's mortality experience is identical to that of IT.
- The hedger's annuity portfolio is large enough so that diversifiable risk can be ignored.
- The hedger establishes a delta-neutral longevity hedge with a freshly launched S-forward at time $t_0 = 2018$. No adjustment is made to the hedge after time t_0 .
- When calibrating the hedge, the hedger assumes that $\{K_t\}$ follows a random walk with drift, and assumes $\{k_t^{(EW)}\}$ and $\{k_t^{(IT)}\}$ follow either ARFIMA or ARMA

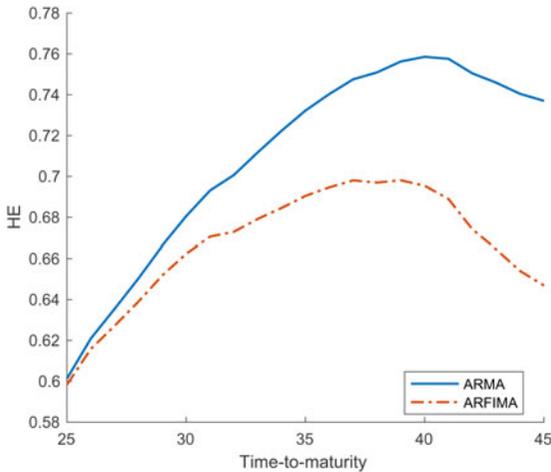


Figure 8. The values of HE for a delta-neutral hedge constructed using a S-forward with a time-to-maturity ranging from 25 to 45 years, when $\{k_t^{(EW)}\}$ and $\{k_t^{(IT)}\}$ are modeled by ARMA (solid line) and ARFIMA (dot-dashed line) processes.

processes. Using ARMA processes means that the hedger ignores or is not aware of the empirical fact that there exists long memory in mortality differentials.

- At time $t_0 = 2018$, freshly launched S-forwards with a reference age 40 and times-to-maturity up to 50 years are available. The reference population of these S-forwards is EW. Therefore, the hedge is subject to population basis risk.
- The hedge effectiveness is gauged using HE as defined in Equation (12), with $i = IT$, and $u = u_{\Delta}^{(IT,EW)}$. The evaluation model (from which realizations of $\mathcal{L}^{(IT)} - u_{\Delta}^{(IT,EW)} \mathcal{H}^{(EW)}$ and $\mathcal{L}^{(IT)}$ given \mathcal{F}_{t_0} are simulated) is assumed to be the same as the one used for calibrating the hedge. In other words, if the hedger overlooks long memory when calibrating the hedge, he/she also ignores long memory when evaluating the hedge.
- When discounting future cash flows, a constant interest rate of $r = 2\%$ per annum is used for all durations.

Figure 8 reports the resulting values of HE for S-forward times-to-maturity ranging from 25 to 45 years. Regardless of whether long memory is taken into account, the value of HE is the highest when the time-to-maturity of the S-forward is around 40 years.

For any S-forward time-to-maturity, the value of HE when long memory is ignored (ARMA processes are used) is higher than that when long memory is taken into account (ARFIMA processes are used). When the time-to-maturity exceeds 40 years, the difference is more than 10 percentage points. The results presented in Figure 8 clearly point to the conclusion that ignoring long memory in mortality differentials would lead to an over-estimation of hedge effectiveness.

To understand the reason behind the phenomenon observed in Figure 8, let us analyze the constituents of the longevity risk faced by the hedger. Under our modeling framework, the hedger faces hedgeable risk that arises from the uncertainty

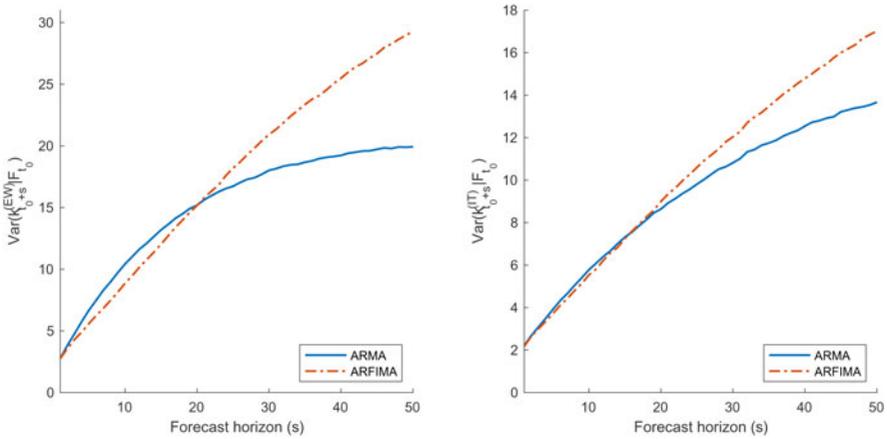


Figure 9. The variance of $k_{t_0+s}^{(EW)}$ (left panel) and $k_{t_0+s}^{(IT)}$ (right panel) given \mathcal{F}_{t_0} for $s=1, \dots, 50$ when $\{k_t^{(EW)}\}$ and $\{k_t^{(IT)}\}$ are respectively modeled by an ARMA process (solid line) and an ARFIMA process (dot-dashed line).

surrounding K_t for $t > t_0$ and unhedgeable risk (population basis risk) that arises from the uncertainty surrounding both $k_t^{(IT)}$ and $k_t^{(EW)}$ for $t > t_0$. Figure 9 shows the variances of $k_{t_0+s}^{(IT)}$ and $k_{t_0+s}^{(EW)}$ given \mathcal{F}_{t_0} for $s=1, \dots, 50$, implied by the ARMA and ARFIMA processes. In line with the results presented in section 2, as the forecast horizon s increases, the ARFIMA processes yield larger variances of $k_{t_0+s}^{(IT)}$ and $k_{t_0+s}^{(EW)}$ given \mathcal{F}_{t_0} compared to the ARMA processes. More specifically, the variances of $k_{t_0+s}^{(IT)}$ and $k_{t_0+s}^{(EW)}$ given \mathcal{F}_{t_0} implied by the ARMA processes (which incorporate short-term memory only) converge to their respective constant levels fairly quickly, while those implied by the ARFIMA processes (which incorporate both short- and long-term memory) grow at slowly decreasing rates. In effect, when $\{k_t^{(IT)}\}$ and $\{k_t^{(EW)}\}$ are modeled by ARFIMA processes (which incorporate the empirical fact that long memory exists in $\{k_t^{(IT)}\}$ and $\{k_t^{(EW)}\}$) instead of ARMA processes, the proportion of hedgeable risk relative to total risk becomes smaller and so does hedge effectiveness.

4.2. Case study II

Case study II is based on the same assumptions as those made in case study I, except the following:

- The annuitant’s mortality experience is identical to that of NL. The hedge is still subject to population basis risk, as the reference population of the S-forward is EW.
- When calibrating the hedge, the hedger assumes that $\{K_t\}$ follows a random walk with drift, and $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$ may follow either ARFIMA or ARMA processes depending on whether long memory is ignored. The hedge may use a delta-neutral hedge or variance-minimizing hedge. These options produce four hedging scenarios in total.
- We use HE defined in Equation (12) to evaluate the performance of each hedge. For all four hedging scenarios, the evaluation model is based on ARFIMA

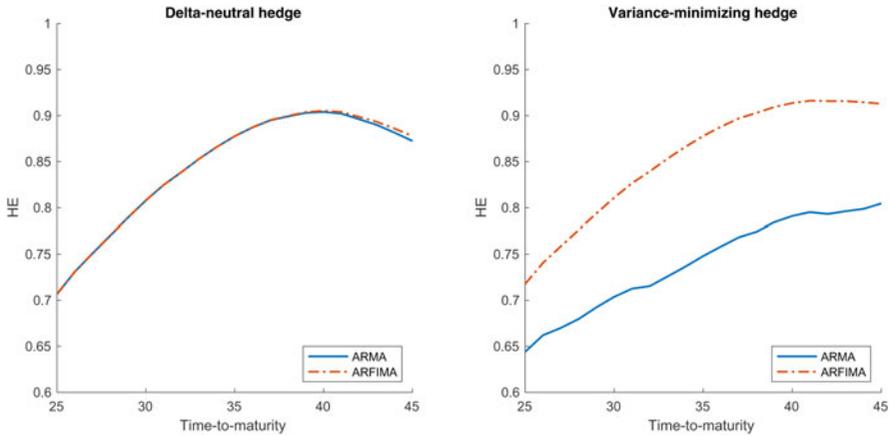


Figure 10. Values of HE produced by delta-neutral hedges (left panel) and variance-minimizing hedges (right panel), for S-forward times-to-maturity ranging from 25 to 45 years, when $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$ are modeled by ARMA processes (solid lines) and ARFIMA processes (dot-dashed lines).

processes for $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$. In other words, we are assuming that long memory exists in reality, but the hedger may or may not incorporate this property when calibrating the hedge.

Our goal is to investigate how delta-neutral and variance-minimizing hedges may underperform if long memory exists in reality but is not accounted for in the hedges.

Figure 10 shows the resulting values of HE for S-forward times-to-maturity ranging from 25 to 45 years. For the variance-minimizing hedge, hedge ratios calculated using the ARMA model assumption lead to materially smaller HE values compared to those computed using the ARFIMA model assumption. However, interestingly, for the delta-neutral hedge, the values of HE produced by hedge ratios computed using the ARMA and ARFIMA model assumptions are highly similar. These results suggest that in reality when long memory in mortality differentials exists, calculating hedge ratios using short memory processes like ARMA would lead to a significant underperformance if the calibration method is variance minimization, but not if the calibration method is delta neutralization.

To obtain further insights, let us examine the hedge ratios calculated in each of the four hedging scenarios (Figure 11). For the variance-minimizing hedge, hedge ratios computed using the ARMA model assumption are consistently smaller than those calculated using the ARFIMA model assumption; as a smaller than optimal notional amount of S-forward is used, the ARMA model assumption leads to underperformance. For the delta-neutral hedge, the hedge ratios are almost unaffected by the model assumption for $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$; hence, the hedge does not underperform even when the ARMA model assumption is used.

We now explain why the hedge ratio for the variance-minimizing hedge is sensitive to the model assumption for $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$, but that for the delta-neutral hedge is not. For the variance-minimizing hedge, the hedge ratio is calculated using Equation

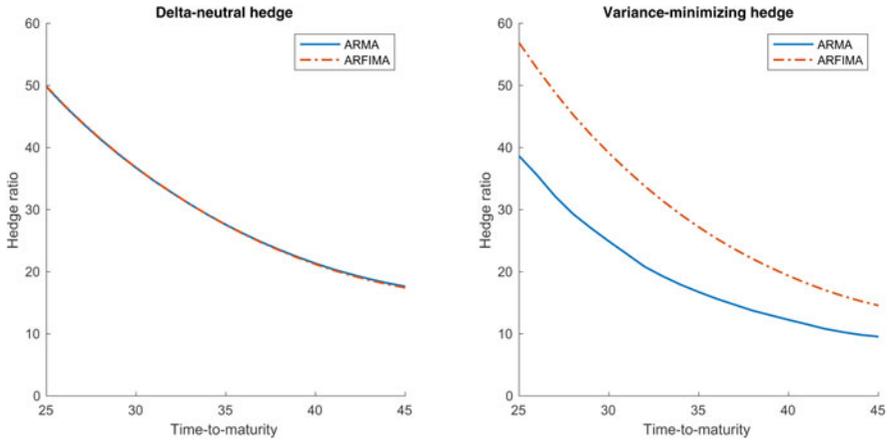


Figure 11. Hedge ratios for delta-neutral hedges (left panel) and variance-minimizing hedges (right panel), for S-forward times-to-maturity ranging from 25 to 45 years, when $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$ are modeled by ARMA processes (solid lines) and ARFIMA processes (dot-dashed lines).

(11), of which both the numerator and denominator are highly sensitive to the model assumption because, as argued earlier, it affects the mix of hedgeable risk (the uncertainty surrounding K_t for $t > t_0$) and unhedgeable risk (the uncertainty surrounding $k_t^{(EW)}$ and $k_t^{(NL)}$ for $t > t_0$). For the delta-neutral hedge, the hedge ratio is obtained from Equation (10), which can be expanded to the following:

$$u_{\Delta}^{(NL,EW)} = \frac{\Delta_L^{(NL)}}{\Delta_H^{(EW)}} = \frac{\frac{\partial}{\partial K_{t_0}} E[\mathcal{L}^{(NL)} | \mathcal{F}_{t_0}]}{\frac{\partial}{\partial K_{t_0}} E[\mathcal{H}^{(EW)} | \mathcal{F}_{t_0}]}.$$

Although the equation above involves $k_t^{(EW)}$ and $k_t^{(NL)}$ for $t > t_0$ through the two expectations, it depends a lot more heavily on the rate of change of the two expectations relative to changes in K_{t_0} . As a result, $u_{\Delta}^{(NL,EW)}$ is quite insensitive to the model assumption for $\{k_t^{(EW)}\}$ and $\{k_t^{(NL)}\}$.

5. Concluding remarks

In this paper, we found empirically that long memory exists in mortality differentials. To capture this empirical fact, we propose modeling the population-specific period effects ($k_t^{(i)}$ for $i \in \mathbb{P}$) in the LL model with ARFIMA processes instead of ARMA processes, as ARFIMA processes are capable of taking both long- and short-term memory into account.

Incorporation of long memory using ARFIMA processes results in mortality forecasts with the following properties. First, the forecasts are coherent in the sense that the expected mortality trajectories between any two populations being modeled do not diverge indefinitely. Second, the rate of convergence to the long-run equilibrium becomes slower. This property may be considered as desirable, because, as Li *et al.* (2017) mentioned, given the patterns of the population-specific period

effects over the sample period, “it does not seem straightforward to justify immediate, quick convergence to the long-term equilibrium.” Third, the forecast uncertainty for $k_t^{(i)}$ for $i \in \mathbb{P}$ over the long run is higher, a property that may guide practitioners to making more adequate provision for population basis risk.

One limitation of our contribution is its data requirement. A sufficiently lengthy data series is required for fitting an ARFIMA process. In our empirical work where a sample period of 1900–2018 is used, the fractional difference parameter d in the ARFIMA process is statistically significant. However, a shorter data series, say one that begins in 1950, may not provide sufficient statistical evidence for the significance of parameter d . This data requirement problem is also noted by Hyndman *et al.* (2013), who apply ARFIMA processes to their product-ratio mortality modeling method.

We presented two case studies to demonstrate the impact of long memory in mortality differentials on index-based longevity hedges. The first case study reveals that overlooking long memory when calibrating and evaluating an index-based longevity hedge would lead to an overly optimistic estimate of the effectiveness of the hedge. The second case study points to the conclusion that compared to a variance-minimizing hedge, a delta-neutral hedge is more robust than a variance-minimizing hedge relative to the inclusion/exclusion of long memory in mortality differentials, because its hedge ratio has negligible dependence on the volatility implied by the processes for the population-specific period effects.

We acknowledge that certain features of mortality dynamics are not considered in this paper. For instance, $\{K_t\}$ and/or $\{k_t^{(i)}\}$ for some $i \in \mathbb{P}$ in the LL model may exhibit conditional heteroskedasticity [Zhou and Li (2020)], a feature that can be captured by generalized autoregressive conditional heteroskedasticity (GARCH) processes. By utilizing a combination of GARCH and AFRIMA processes, we may study in future research the interaction between long memory and conditional heteroskedasticity, as well as its impact on index-based longevity hedges.

References

- Cairns, A. J. (2011) Modelling and management of longevity risk: approximations to survivor functions and dynamic hedging. *Insurance: Mathematics and Economics* 49(3), 438–453.
- Cairns, A. J., D. Blake, K. Dowd, G. D. Coughlan and M. Khalaf-Allah (2011) Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin: The Journal of the IAA* 41(1), 29–59.
- Cairns, A. J., K. Dowd, D. Blake and G. D. Coughlan (2014) Longevity hedge effectiveness: a decomposition. *Quantitative Finance* 14(2), 217–235.
- Coughlan, G. D., M. Khalaf-Allah, Y. Ye, S. Kumar, A. J. Cairns, D. Blake and K. Dowd (2011) Longevity hedging 101: a framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal* 15(2), 150–176.
- Dahl, M., M. Melchior and T. Møller (2008) On systematic mortality risk and risk-minimization with survivor swaps. *Scandinavian Actuarial Journal* 2008(2–3), 114–146.
- Delgado-Vences, F. and A. Ornelas (2019) Modelling Italian mortality rates with a geometric-type fractional Ornstein-Uhlenbeck process. preprint [arXiv:1901.00795](https://arxiv.org/abs/1901.00795).
- Dowd, K., A. J. Cairns, D. Blake, G. D. Coughlan and M. Khalaf-Allah (2011) A gravity model of mortality rates for two related populations. *North American Actuarial Journal* 15(2), 334–356.
- Enchev, V., T. Kleinow and A. J. Cairns (2017) Multi-population mortality models: fitting, forecasting and comparisons. *Scandinavian Actuarial Journal* 2017(4), 319–342.
- Gil-Alana, L. A., J. Cunado and R. Gupta (2017) Persistence, mean-reversion and non-linearities in infant mortality rates. *Social Indicators Research* 131(1), 393–405.
- Granger, C. W. and R. Joyeux (1980) An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis* 1(1), 15–29.

- Hyndman, R. J., H. Booth and F. Yasmeen (2013) Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography* 50(1), 261–283.
- Kleinow, T. (2015) A common age effect model for the mortality of multiple populations. *Insurance: Mathematics and Economics* 63, 147–152.
- Li, J. S.-H., W. -S. Chan and R. Zhou (2017) Semicoherent multipopulation mortality modeling: the impact on longevity risk securitization. *Journal of Risk and Insurance* 84(3), 1025–1065.
- Li, J. S.-H. and M. R. Hardy (2011) Measuring basis risk in longevity hedges. *North American Actuarial Journal* 15(2), 177–200.
- Li, J. S.-H., J. Li, U. Balasooriya and K. Q. Zhou (2021) Constructing out-of-the-money longevity hedges using parametric mortality indexes. *North American Actuarial Journal* 25(sup1), S341–S372.
- Li, J. S.-H., R. Zhou and M. Hardy (2015) A step-by-step guide to building two-population stochastic mortality models. *Insurance: Mathematics and Economics* 63, 121–134.
- Li, N. and R. Lee (2005) Coherent mortality forecasts for a group of populations: an extension of the Lee-Carter method. *Demography* 42(3), 575–594.
- Lo, A. W. (1991) Long-term memory in stock market prices. *Econometrica: Journal of the Econometric Society* 59, 1279–1313.
- Villegas, A. M., S. Haberman, V. K. Kaishev and P. Millossovich (2017) A comparative study of two-population models for the assessment of basis risk in longevity hedges. *ASTIN Bulletin: The Journal of the IAA* 47(3), 631–679.
- Wang, L., M. C. Chiu and H. Y. Wong (2021) Volterra mortality model: actuarial valuation and risk management with long-range dependence. *Insurance: Mathematics and Economics* 96, 1–14.
- Wang, L. and H. Y. Wong (2021) Time-consistent longevity hedging with long-range dependence. *Insurance: Mathematics and Economics* 99, 25–41.
- Yan, H., G. W. Peters and J. S. Chan (2020) Multivariate long-memory cohort mortality models. *ASTIN Bulletin: The Journal of the IAA* 50(1), 223–263.
- Yan, H., G. W. Peters and J. Chan (2021) Mortality models incorporating long memory for life table estimation: a comprehensive analysis. *Annals of Actuarial Science* 151–38.
- Yaya, O. S., L. A. Gil-Alana and A. Y. Amoateng (2019) Under-5 mortality rates in G7 countries: analysis of fractional persistence, structural breaks and nonlinear time trends. *European Journal of Population* 35(4), 675–694.
- Zhou, K. Q. and J. S.-H. Li (2017) Dynamic longevity hedging in the presence of population basis risk: a feasibility analysis from technical and economic perspectives. *Journal of Risk and Insurance* 84(S1), 417–437.
- Zhou, K. Q. and J. S.-H. Li (2020) Asymmetry in mortality volatility and its implications on index-based longevity hedging. *Annals of Actuarial Science* 14(2), 278–301.
- Zhou, K. Q. and J. S.-H. Li (2021) Longevity Greeks: what do insurers and capital market investors need to know?. *North American Actuarial Journal* 25(sup1), S66–S96.
- Zhou, R., Y. Wang, K. Kaufhold, J. S.-H. Li and K. S. Tan (2014) Modeling period effects in multi-population mortality models: applications to Solvency II. *North American Actuarial Journal* 18(1), 150–167.