ON THE NATURE AND HANDLING OF COVID-19 IN ARIMA MORTALITY FORECASTS

By S. J. Richards

LONGEVITAS™

On the nature and handling of COVID-19 in ARIMA mortality forecasts

Richards, Stephen J.*

June 29, 2026

Abstract

The COVID-19 pandemic caused outliers in many nations' mortality data in 2020. Where such outliers are present, they can severely affect the calibration of mortality-projection models. If not handled correctly, such outliers can lead to biased mortality forecasts or excessive insurer capital requirements for longevity trend risk. We use the methodology of Chen and Liu [1993] to investigate the nature of the outlier caused by COVID-19 in four European countries, and to create unbiased ARIMA forecasts with the Lee-Carter model. We present evidence that COVID-19 represented (i) no outlier among females in Sweden, (ii) a level-shift outlier in the Netherlands, and (iii) a multi-year temporary change in Italy, the UK and for males in Sweden. The mortality effects of the COVID-19 pandemic thus appear to be persistent to varying degrees after 2020. We show how continuing elevated post-pandemic mortality can be reflected by splitting an ARIMA(p, 1, q) mortality forecast into its three constituent components: (i) a linear trend, (ii) an ongoing outlier effect, and (iii) a cumulative stochastic deviation. Finally, we show that there is a risk of later reclassification of outliers, with consequences for actuaries reserving for pensions and annuities.

Keywords: COVID-19, mortality shocks, outliers, robust forecasting, ARIMA.

1 Introduction

Mortality projections are an important part of demographic work, including population projections, forecast life expectancy, and forecast dependency ratios. Mortality projections are also an important part of actuarial work, including reserving for pensions and annuities or the pricing of bulk annuities and longevity reinsurance. Several important stochastic mortality-projection models rely on forecasting a univariate time index; examples include the model from Lee and Carter [1992] and the Age-Period-Cohort (APC) model. Such univariate mortality time indices are often well represented by an ARIMA model [Kleinow and Richards, 2016]. In this paper we will focus on actuarial aspects of mortality forecasting with ARIMA models, but many of the points will also apply to non-actuarial demographic work.

However, ARIMA models are sensitive to outliers [Martin et al., 1983, p.2]. An outlier is an observation that materially distorts the parameterisation of an ARIMA model such that it leads to (i) an incorrect central mortality forecast, (ii) a false statement of uncertainty around the forecast, or (iii) a false statement about the capital to held in respect of longevity trend risk. For accurate mortality forecasting we need ARIMA models that are not distorted by outliers. As demonstrated in Figure 1, failure to handle outliers produces a poor-quality forecast.

*stephen@longevitas.co.uk, Longevitas Ltd, Edinburgh. www.longevitas.co.uk. http://orcid.org/0000-0001-6859-6347.

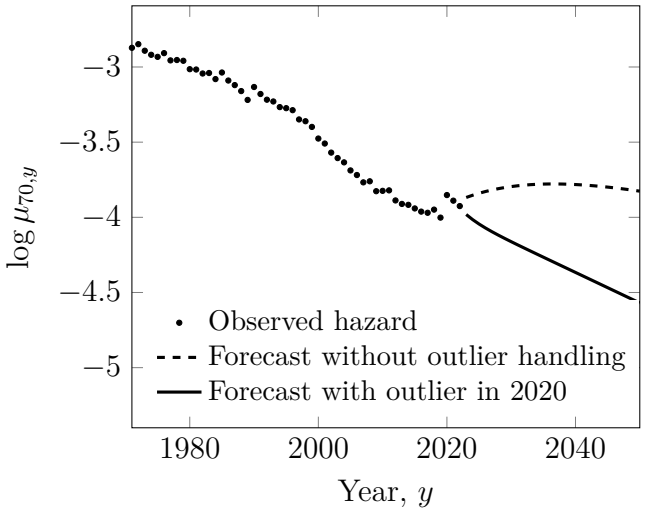

Figure 1: Logarithm of mortality hazard at age 70 in year y for males in UK — observed hazard (•) and forecast hazard using ARIMA(p, 1, q) models for κ_y in a Lee-Carter model with and without outlier handling. Source: own calculations using HMD data 1971–2022. The outlier is a temporary change (TC), as described Section 4.

The mortality shock in 2020 associated with the COVID-19 pandemic [The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020] provided an outlier in many countries' mortality data. Richards [2024] considered outlier-detection approaches for a wide variety of stochastic mortality models. In each case, models were calibrated to UK mortality data up to and including 2020, i.e. the outlier lay at the end of data period and thus at the end of any mortality index. In the case of ARIMA models for forecasting, Richards [2024, Section 4] provisionally assumed that 2020 was an additive outlier (AO). At the time this assumption was of little consequence, as "it is impossible to empirically distinguish the type of an outlier occurring at the very end of a series" [Chen and Liu, 1993, page 286].

The availability of subsequent post-pandemic data permits us to address further questions, including what type of outlier was COVID-19? Is the outlier type the same for each country? Is the outlier type the same for both sexes within a country? This paper seeks to answer these questions as follows: Section 2 describes the data used; Section 3 describes how the period effects are estimated; Section 4 describes the outlier types available in Chen and Liu [1993], and which are of relevance to actuaries; Section 5 describes the structure of a regression ARIMA model without outliers; Section 6 extends this to show the structure of a regression ARIMA model allowing for outliers; Section 7 looks at the nature of COVID-19 outliers in four European countries, while Section 8 considers the risk of later outlier reclassification. Section 9 concludes.

Throughout this paper we will use κ_y to denote the univariate mortality period effect in calendar year y. We follow Harvey [1981, equation (1.9)] in using Δ to denote the backward difference operator, i.e. $\Delta\kappa_y = \kappa_y - \kappa_{y-1}$. Following Shumway and Stoffer [2010, page 61] we will also use B to denote the backshift operator, i.e. $B\kappa_y = \kappa_{y-1}$ and $B^i\kappa_y = \kappa_{y-i}$. We will sometimes use polynomials in B for conciseness and note that $\Delta\kappa_y = (1 - B)\kappa_y$. In other publications the backshift operator, B, is referred to as the lag operator, L, but the two are synonymous. An estimate of κ_y will be denoted $\hat{\kappa}_y$, and estimation will typically be by maximum likelihood.

2 Data

We use population data from the Human Mortality Database for lives aged 50 and over, as this age range is most relevant for pensions and annuities in payment. We use data for Italy (1971–2023), the Netherlands (1971-2023), Sweden (1971–2024) and the UK (1971–2022). In each case we have the numbers of lives observed dying at age x in year y, $d_{x,y}$, and the corresponding mid-year exposed-to-risk, $E^c_{x,y}$. The data lend themselves naturally to fitting a model for the mortality hazard, $\mu_{x+0.5,y+0.5}$, but we drop the 0.5 henceforth for simplicity. The death counts, $d_{x,y}$, are realisations of a pseudo-Poisson count, $D_{x,y}$; as per Macdonald and Richards [2025], 'pseudo-Poisson' means that the log-likelihood is identical to that of a Poisson model apart from normalising constants that don't affect maximum-likelihood estimation. We can therefore treat the death counts, $d_{x,y}$, as though they are realisations of Poisson random variables, although it is useful to remember that they are not truly Poisson. In addition, $d_{x,y}$ will exhibit over-dispersion in practice [Djeundje and Currie, 2011], although this tends only to affect variance estimates and smoothing levels. Since this paper is primarily concerned with mean estimation, we will ignore over-dispersion for simplicity. Plots of d_x at the start and end of each data set are shown in Figure 2.

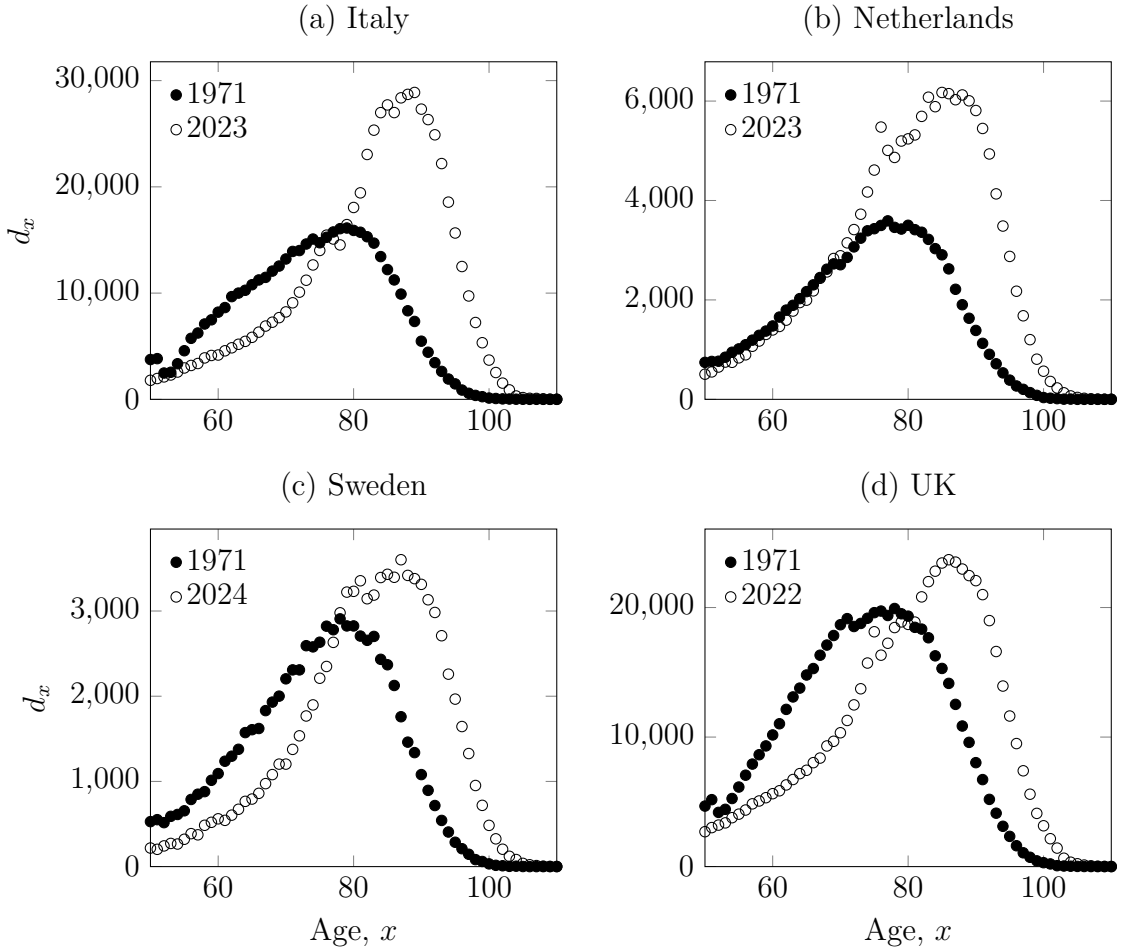

Figure 2: Deaths by age, d_x, for four European countries at the start and end of each data set (males and females combined). Source: own calculations using data from the Human Mortality Database, accessed 12th April 2026.

3 Estimating period effects

Period effects are denoted κ_y in the Lee-Carter [Lee and Carter, 1992] model for the mortality hazard at age x and year y:

$$\log \mu_{x,y} = \alpha_x + \beta_x \kappa_y \tag{1}$$

which is fitted using the identifiability constraints described in Appendix A. We follow Brouhns et al. [2002] and estimate parameters by maximising the following log-likelihood:

$$\ell = \sum_x \sum_y \left(d_{x,y} \log \mu_{x,y} - E^c_{x,y} \mu_{x,y} \right) \tag{2}$$

Maximum-likelihood estimation is preferable to the singular value decomposition (SVD) approach of Lee and Carter [1992] or Girosi and King [2008], because SVD does not account for the widely differing variances of $\log \mu_{x,y}$ over age and time. We use the algorithm of Currie [2013] to apply penalised-spline smoothing to α_x and β_x during parameter estimation, i.e. simultaneous estimation and smoothing. Smoothing not only reduces the dimensionality of the model, but smoothing of β_x in particular improves forecast quality by minimising the crossover of mortality rates at adjacent ages [Delwarde et al., 2007]. We do not smooth κ_y because of our prior expectation that there is not a smooth underlying process. This differs from Richards and Currie [2009], who did smooth the κ_y terms in the Lee-Carter model. However, in practice attempting to smooth the κ_y terms can lead to under-smoothing - and volatile forecasts, even without pandemic shocks.

Regarding equation (1), Lee and Carter [1992, Appendix B] argued that uncertainty over $\hat{\kappa}$ dominates the forecast. We therefore make the simplifying assumption that $\text{Var}(\log \hat{\mu}_{x,y}) = \hat{\beta}^2_x \text{Var}(\hat{\kappa}_y)$. For forecasting purposes we will treat the $\hat{\kappa}_y$ estimates as being observations of an ARIMA(p, d, q) process, where p is the number of autoregressive parameters, d is the order of differencing and q is the number of moving-average terms; see Gardner et al. [1980] or Shumway and Stoffer [2010, Section 3] for an introduction to ARIMA models for time series. Due to the close correspondence between mortality-improvement rates and $\Delta\hat{\kappa}_y$ (see Appendix B), we will limit our attention to ARIMA(p, 1, q) models. Of course, the $\hat{\kappa}_y$ values are parameter estimates, not observed data, so this represents a further simplifying assumption. In particular, the true underlying variance of $\hat{\kappa}_y$ will likely be heteroscedastic as population sizes have grown over time and the distribution of age at death has changed shape (Figure 2).

The Lee-Carter models are fitted and the resulting estimates $\hat{\kappa}_y$ are plotted in Figures 3-6 for four European countries. We note that most of the $\Delta\hat{\kappa}_y$ values in Figures 3-6 lie below zero, i.e. we require a non-zero mean or drift term, μ, in our ARIMA models for κ_y. Various outliers are evident for 2020, although for females in Sweden (Figure 5b) it is unclear if there is any outlier at all. Sometimes it is easier to spot the outlier in the undifferenced $\hat{\kappa}_y$ values, other times it is clearer to look at the differences, $\Delta\hat{\kappa}_y$. Of particular interest is that COVID-19 appears to be a modest event for Sweden in Figure 5, whereas for other countries COVID-19 led to a multi-year elevated mortality that decayed over subsequent years. In this paper we will use the methodology of Chen and Liu [1993] to formally test for different types of outlier, and in Section 6 we will consider how multi-year outlier effects are carried forward into the ARIMA forecast of κ_y.

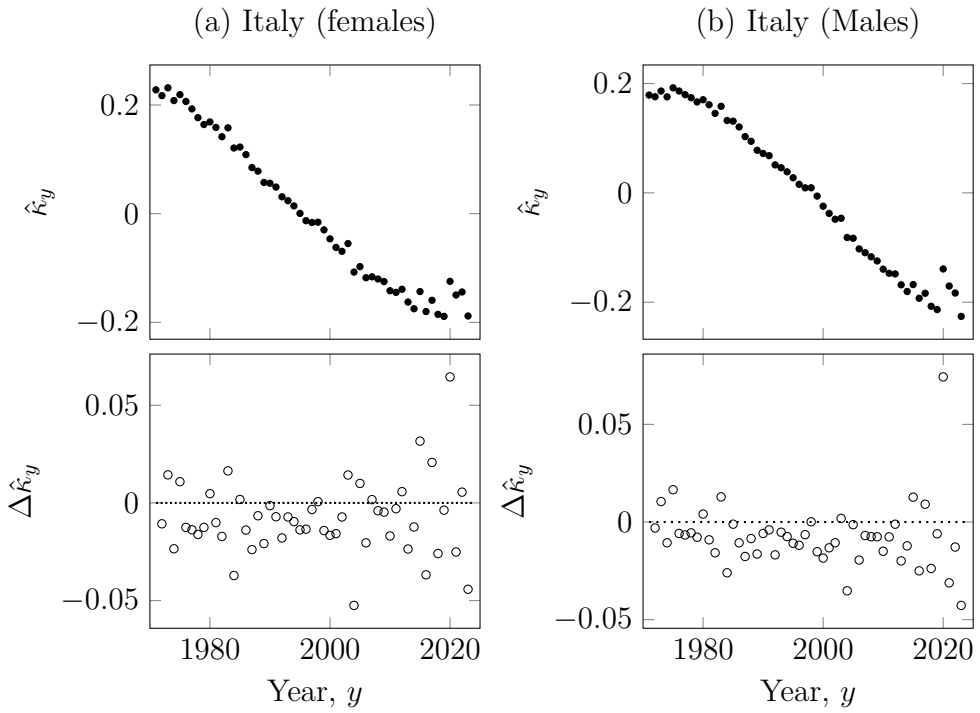

Figure 3: Estimated period effects for Lee-Carter model in equation (1) calibrated using data for lives in Italy aged 50 and over. Top row: $\hat{\kappa}_y$; bottom row: $\Delta\hat{\kappa}_y$. Source: own calculations using data from the Human Mortality Database, accessed 12th April 2026.

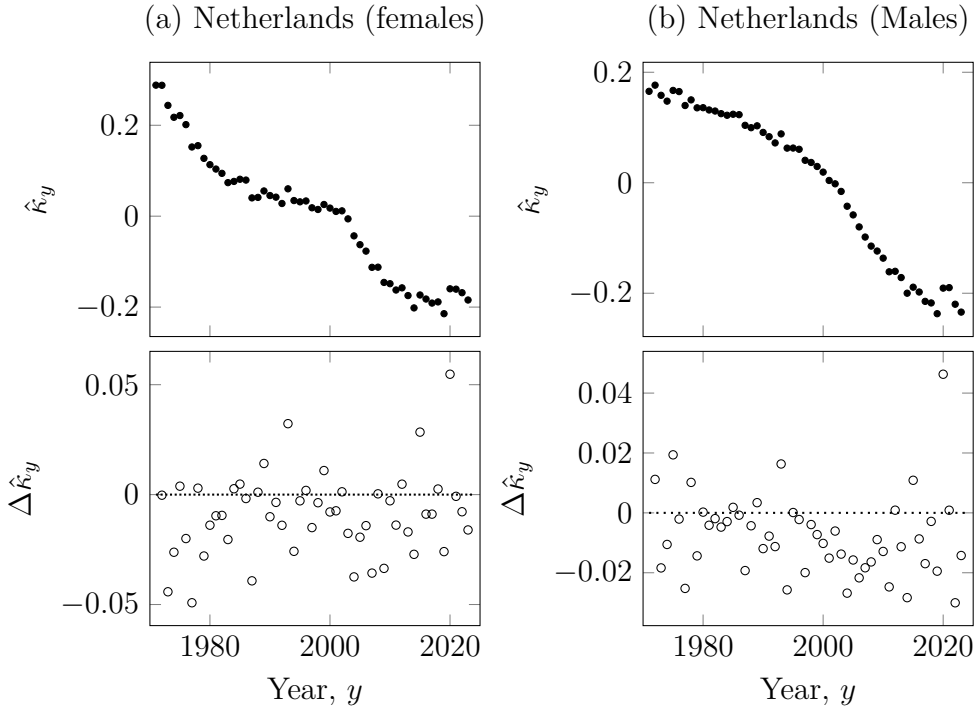

Figure 4: Estimated period effects for Lee-Carter model in equation (1) calibrated using data for lives in Netherlands aged 50 and over. Top row: $\hat{\kappa}_y$; bottom row: $\Delta\hat{\kappa}_y$. Source: own calculations using data from the Human Mortality Database, accessed 12th April 2026.

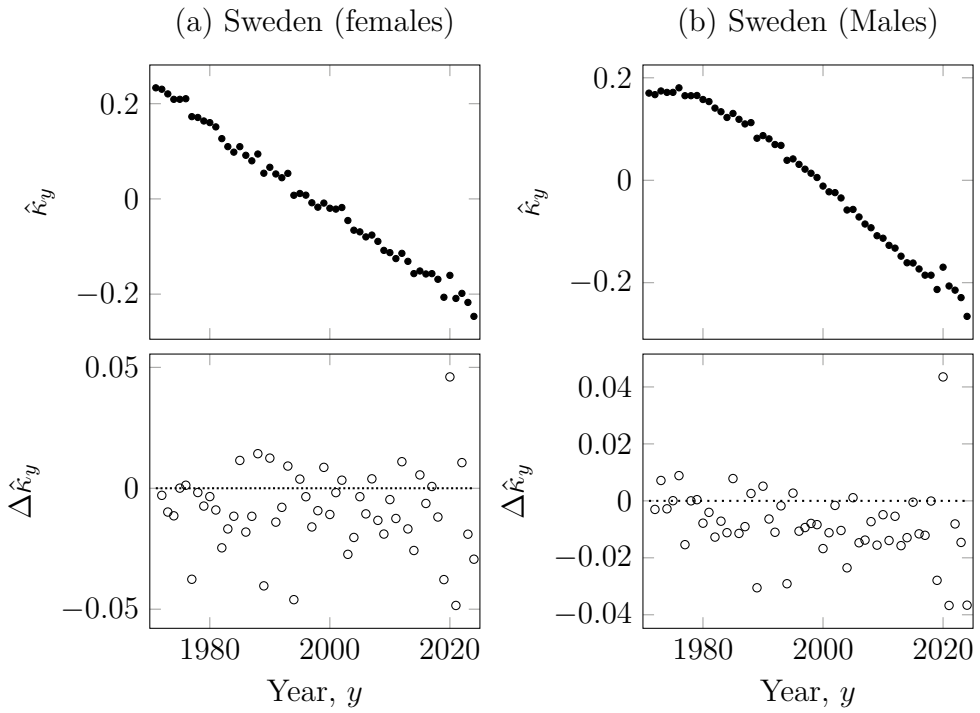

Figure 5: Estimated period effects for Lee-Carter model in equation (1) calibrated using data for lives in Sweden aged 50 and over. Top row: $\hat{\kappa}_y$; bottom row: $\Delta\hat{\kappa}_y$. Source: own calculations using data from the Human Mortality Database, accessed 12th April 2026.

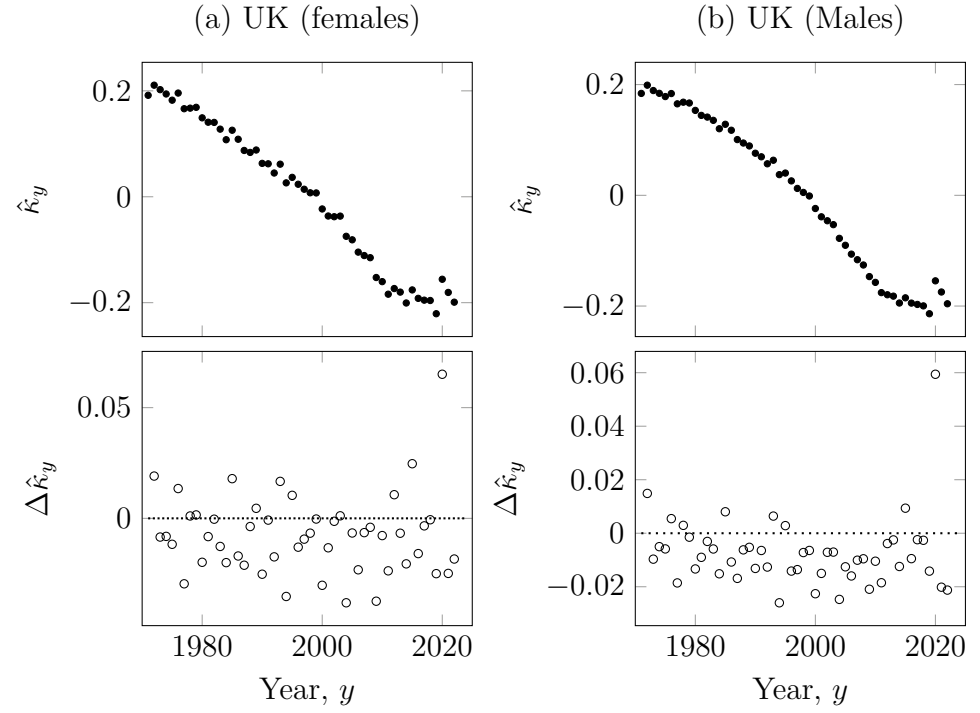

Figure 6: Estimated period effects for Lee-Carter model in equation (1) calibrated using data for lives in UK aged 50 and over. Top row: $\hat{\kappa}_y$; bottom row: $\Delta\hat{\kappa}_y$. Source: own calculations using data from the Human Mortality Database, accessed 12th April 2026.

4 Outlier types

Chen and Liu [1993] describe four types of outlier for ARIMA models: innovation outlier (IO), additive outlier (AO), level shift (LS) and temporary change (TC). Richards [2024, Figure 3] argued that actuaries should always leave IOs unhandled, as allowing for them suppresses the estimated variance of the innovation process, $\hat{\sigma}^2_\epsilon$. Thus, omitting IOs when calibrating ARIMA models for $\hat{\kappa}_y$ leads to unduly low value-at-risk capital requirements [Kleinow and Richards, 2016, Section 7]. We therefore leave in IOs for actuarial work. The question then is whether COVID-19 represents an AO, LS, TC or none of the three?

We begin by plotting some stylised outliers in an undifferenced stationary series without drift, X_t, in Figure 7(a)-(c). The AO in Figure 7(a) represents a one-off mortality event with no downstream consequences. The LS in Figure 7(b) represents a permanent step change. The TC in Figure 7(c) causes an effect that persists over multiple time periods while decaying in strength. The bottom row in Figure 7 shows the corresponding patterns in the first differences.

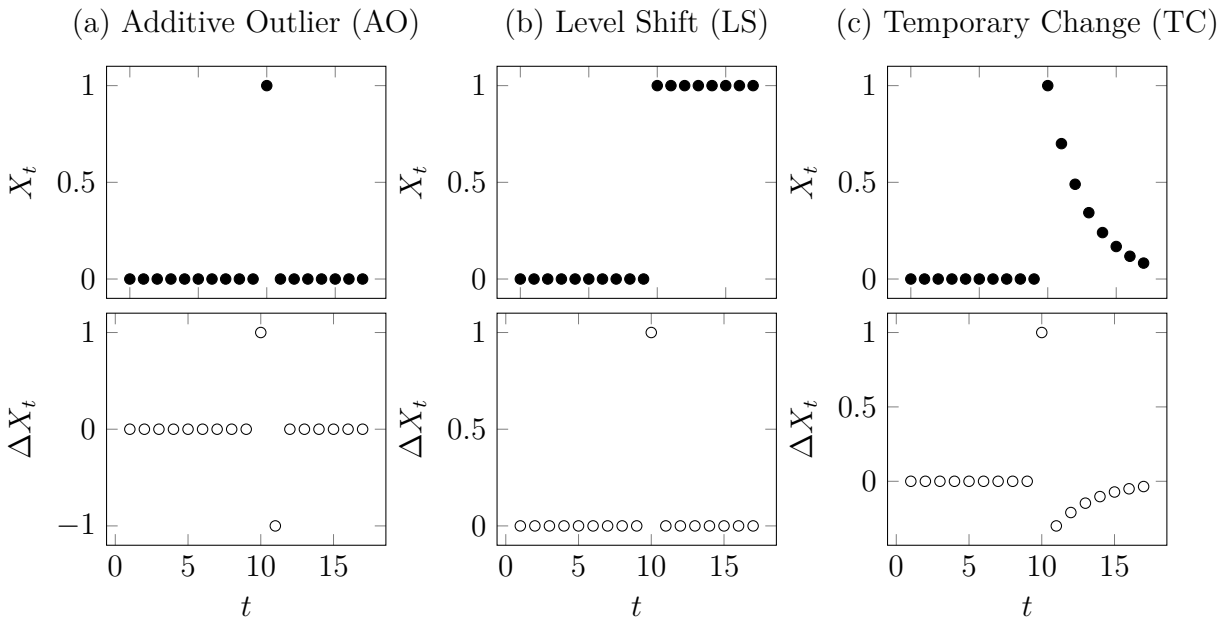

Figure 7: Types of outliers in undifferenced series (•, top row) and their equivalent implied outlier patterns in the differenced series (∘, bottom row). Note the equivalence of an LS in the undifferenced series to an AO in the differenced series.

Looking at Figure 6, it seems that identifying the location of the outlier is easiest in $\Delta\hat{\kappa}_y$, but that outlier classification is visually easier in the $\hat{\kappa}_y$ values. In particular, it looks like COVID-19 caused a TC outlier for males and females in both Italy and the UK. Visual evidence further suggests that COVID-19 might have caused an LS outlier among females in the Netherlands. However, visual assessment is inadequate — we will formally assess the outlier types in Section 7 and consider the risk of outlier reclassification in Section 8.

Note that this paper is concerned with outlier identification and classification for ARIMA models. There are other approaches to outliers caused by COVID-19, such as Cairns and Blake [2026]. There is also a direct correspondence between the outlier types in Figure 7 and the jump processes of Goes et al. [2025]: an AO is a transitory jump, an LS is a permanent jump and a TC is a vanishing jump.

5 ARIMA model structure without outliers

Let y be the last year of the data and let $y+t$ denote each year in the forecast period, $t \in 1, 2, \ldots$. Let κ_{y+t} denote the univariate mortality index in year $y+t$. We assume that κ_{y+t} follows a first-order ARIMA process with p autoregressive (AR) terms and q moving-average (MA) terms as follows:

$$\phi(B)(1-B)\kappa_{y+t} = \mu + \theta(B)\epsilon_{y+t} \tag{3}$$

where μ is the mean or drift term and ϵ_{y+t} is an innovation term with mean zero, constant variance σ_ϵ^2 and where $\text{Cov}(\epsilon_i, \epsilon_j) = 0, \forall i \neq j$. $\phi(B)$ and $\theta(B)$ are polynomials in the backshift operator, B:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \text{ for the AR terms, and} \tag{4}$$
$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q \text{ for the MA terms.} \tag{5}$$

In this paper we restrict our attention to models with $p, q \in \{0, 1, 2, 3\}$. Equations (4) and (5) use the same representation of AR and MA polynomials as Harvey [1981, Section 4], Shumway and Stoffer [2010, page 85] and in R's `arima()` command, although some authors use different representations. Richards [2024, Appendix B] showed that an ARIMA(p, 1, q) forecast for κ_{y+t} is then:

$$\kappa_{y+t} = \underbrace{\kappa_y + t\mu}_{\substack{\text{linear} \\ \text{trend}}} + \underbrace{\sum_{i=1}^{t} \frac{\theta(B)}{\phi(B)}\epsilon_{y+i}}_{\substack{\text{cumulative} \\ \text{stochastic} \\ \text{deviation}}} \tag{6}$$

Equation (6) shows that an ARIMA(p, 1, q) model for κ_{y+t} is equivalent to a linear trend with a cumulative stochastic deviation around that trend. Richards [2024, Appendix B] showed that the same model can be fitted either as an ARIMA model with a mean [Richards, 2024, Figure 14] or else as a regression ARIMA (REGARIMA) model [Richards, 2024, Figure 16]. The latter approach becomes particularly convenient when outliers need to be accounted for.
An alternative way of expressing equation (3) is as an ARMA(p, q) process for $\Delta\kappa_{y+t}$:

$$\phi(B)\Delta\kappa_{y+t} = \mu + \theta(B)\epsilon_{y+t} \tag{7}$$

We can express equation (7) in terms of the ARMA process, Y_t, in Chen and Liu [1993, equation (1)]:

$$Y_t = \frac{\theta(B)}{\phi(B)(1-B)}\epsilon_t \tag{8}$$

where $Y_t = \kappa_{y+t} - t\mu$ with $Y_0 = \kappa_y$. Chen and Liu [1993, Section 1] consider a generalised differencing polynomial $\alpha(B)$, but we follow Kleinow and Richards [2016] by fixing $\alpha(B) = 1 - B$ so that we effectively have a model for mortality improvements (see Appendix B).

For the cumulative stochastic component in equation (6) we have the following:

$$\sum_{i=1}^{t} \frac{\theta(B)}{\phi(B)}\epsilon_{y+i} = \begin{cases} \sum_{i=1}^{t} \theta(B)\epsilon_{y+i} & \text{for models without an AR component} \\ \sum_{i=1}^{t} \psi(B)\epsilon_{y+i} & \text{for models with an AR component} \end{cases} \tag{9}$$

where $\psi(B)$ is the infinite moving-average (MA(∞)) representation of the ARMA(p, q) process, i.e. the solution to $\psi(B)\phi(B) = \theta(B)$. The coefficients of $\psi(B)$ are given by Harvey [1981, equation (4.6)] and an implemention can be found in the `ARMAtoMA()` function in R. Parameter estimates for selected ARIMA processes for κ are given in Table 1; by coincidence, the AICc [Hurvich and Tsai, 1989] is minimised in each case by an ARIMA(1,1,2) model (see Appendix C for the calculation of an information criterion for ARIMA models).

Parameter estimate	Data period from 1971 to: 2013	2016	2019	2022
$\hat{\phi}_1$	0.936371	0.835046	0.743796	0.933420
$\hat{\theta}_1$	−1.612196	−1.286468	−1.163156	−1.657586
$\hat{\theta}_2$	0.839883	0.696757	0.606703	0.903732
$\hat{\mu}$	−0.009653	−0.007754	−0.008283	−0.003733
$\hat{\sigma}^2_\epsilon$	0.000074	0.000072	0.000062	0.000111

Table 1: Parameter estimates for ARIMA(1,1,2) models fitted to κ values for Lee-Carter models for UK males from 1971 to the year shown. No outliers are assumed, even for the period ending 2022. Note how the inclusion of COVID-19-affected data in the period 2020–2022 has inflated the variance estimate, $\hat{\sigma}^2_\epsilon$, which is the main driver for longevity trend risk capital under a value-at-risk (VaR) regime [Kleinow and Richards, 2016].

Each of the models in Table 1 has an MA(∞) representation with an infinite series of ψ coefficients. Figure 8 shows how these coefficients decay to zero, with the principle driver being the estimate $\hat{\phi}_1$.

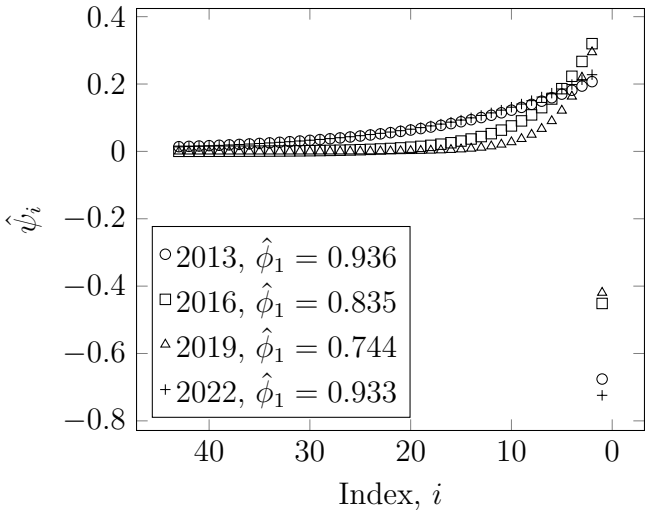

Figure 8: Estimated coefficients, $\hat{\psi}_i$, in MA(∞) representation of the ARMA(1,2) part of the models in Table 1. Values of $\hat{\phi}_1$ closer to 1 result in a slower decay of ψ_i to zero, which means longer-lasting deviations from the linear trend. Note the reversed horizontal axis to show the decaying coefficients applied to error terms in the more distant past towards the left.

Note that the only reliable hallmarks of the outlier in 2020 in Table 1 are the small absolute value $\hat{\mu}$ and the large estimate of $\hat{\sigma}^2_\epsilon$.

Forecasting using equation (6) is handled by R's `predict.Arima()` function, which can be called by passing the object returned by `arima()` to the `predict()` function. `predict.Arima()` manages

the stochastic component of the forecast by using zero values for unobserved future innovations, thus facilitating a gradual transition to the underlying linear forecast in the trend component. The speed of this transition depends on the rate of decay to zero of the ψ coefficients shown in Figure 8, with lower estimates of $\hat{\phi}_1$ leading to faster decay to an ultimately linear forecast.

Simulated sample paths using equation (6) can be generated by taking the linear forecast and generating N(0,σ_ϵ^2) variates for future innovations instead of using zero for a central forecast in `predict.Arima()`. This sample-path simulation is done in R with the `arima.sim()` function. As can be seen in Table 1, a non-robustified ARIMA model will have an inflated estimate $\hat{\sigma}_\epsilon^2$; if this is combined with an estimate of $\hat{\phi}_1$ close to 1, this will lead to wilder simulated sample paths due to slower decay of the ψ_i terms in the MA(∞) representation of the model. Unrobustified ARIMA models will therefore lead to over-statement of longevity trend-risk capital under Solvency II [Kleinow and Richards, 2016]. Using simulated innovations in `arima.sim()` covers the volatility of mortality in the individual forecast years in the stochastic component of equation (6). However, it does not cover the trend uncertainty posed by the uncertainty over $\hat{\mu}$, which can be accounted for by using $\hat{\mu} + Z\text{se}(\hat{\mu})$, where $Z \sim N(0,1)$. The volatility and trend uncertainty can be separately switched on and off to explore their relative contributions to overall forecast or sample-path uncertainty. Kleinow and Richards [2016, Section 7] used this to show that σ_ϵ^2 is the dominant driver of value-at-risk-style capital requirements, not trend uncertainty, at least for Lee-Carter models using ARIMA forecasts of κ_{y+t}.

6 ARIMA model structure with outliers

In equation (6) κ_{y+t} has a linear trend component and a cumulative deviation around the trend. In equation (8) we showed how this could be re-expressed using the notation of Chen and Liu [1993]. We now follow Chen and Liu [1993, equation (19)] and define κ^*_{y+t} as the period effect containing m possible outlier effects:

$$\kappa^*_{y+t} = \kappa_{y+t} + \underbrace{\sum_{j=1}^{m} \omega_j I_j(y+t)}_{\text{outlier component}} \tag{10}$$

where ω_j is the effect of the j^{th} outlier. This gives us an equivalent to equation (6) capable of handling outliers::

$$\kappa^*_{y+t} = \underbrace{\kappa_y + t\mu}_{\substack{\text{linear}\\ \text{trend}}} + \underbrace{\sum_{j=1}^{m} \omega_j I_j(y+t)}_{\text{outlier component}} + \underbrace{\sum_{i=1}^{t} \frac{\theta(B)}{\phi(B)} \epsilon_{y+i}}_{\substack{\text{cumulative}\\ \text{stochastic}\\ \text{deviation}}} \tag{11}$$

For AO and LS outliers $I_j(y+t) \in \{0,1\}$ is an indicator function for the action of the j^{th} outlier in year $y+t$. Outliers occur in the data period, but the effects of some outlier types stretch into the forecast period. Consider an example data period of 1971–2023 with a single outlier in 2020. Table 2 illustrates $I(y+t)$ for the various types of outlier — an AO outlier has no direct effect on the forecast, but LS and TC outliers do. The operation of the TC outlier in Table 2 is reminiscent of the "rocket and feather" description of Bacon [1991], albeit in a very different context — mortality rates are quick to increase, but can take time to fall thereafter. As we will see in Section 7, TC outliers appear to be the commonest response to COVID-19 in the four European countries examined.

Year	AO	LS	TC
1971	0	0	0
1972	0	0	0
⋮	⋮	⋮	⋮
2019	0	0	0
2020	1	1	1
2021	0	1	0.7
2022	0	1	0.49
2023	0	1	0.343
2024	0	1	0.2401
2025	0	1	0.16807
2026	0	1	0.117649

Table 2: Outlier function, $I(y+t)$, for three types of outlier occurring in 2020. The data period covers 1971–2023 (i.e. $y = 2023$), while 2024–2026 are the first three years of the forecast. For the TC outlier we use the default value $\delta = 0.7$ [Chen and Liu, 1993, Section 1.1].

One aspect to note about the function $I_j(y+t)$ in equation (11) and Table 2 is that it combines the roles of $L_j(B)$ and $I_t(t_j)$ in Chen and Liu [1993, equation (19)]. The two approaches have exactly the same effect on the undifferenced series. Indeed, AO and LS outliers can be viewed as special cases of TC outliers by noting that $I_j(y+t)$ for the jth outlier occurring in year y_j can be defined as follows:

$$I_j(y+t) = \begin{cases} 0, & y_j > y+t \\ \delta^{y+t-y_j} & y_j \leq y+t \end{cases} \tag{12}$$

where $\delta = 0$ for an AO, $\delta = 1$ for an LS and $0 < \delta < 1$ for a TC. Equation (12) reproduces the three patterns in Table 2, including AOs, if we follow Knuth [1992, p.6] and IEEE Computer Society [2019, p.63] and adopt the convention that $0^0 = 1$ for integer exponents.

Once the methodology of Chen and Liu [1993] has been used to identify the location, type and effect of outliers, they become part of the outlier component of the model and forecast in equation (11). The same applies to the linear trend $t\mu$ after $\hat{\mu}$ has been estimated. The parameters of the outlier component are estimated from the data period, but for LS and TC outliers their effect continues into the forecast period, as shown in Table 2. The forecast $\{\epsilon_{y+t}\}$ values are different: their value is either set to zero (for a central forecast) or simulated from a $\text{N}(0,\sigma_\epsilon^2)$ distribution (for generating sample paths).

Note that for selecting an ARIMA(p, 1, q) model with outliers we use the AIC with the small-sample correction described in Appendix C, with particular reference to the contribution of outliers to the overall parameter count, k.

7 Nature and timing of outliers in four European countries

Table 3 shows the results of applying the methodology of Chen and Liu [1993] to the mortality data of the four European countries described in Section 2. We fit the Lee-Carter model in equation (1) with smoothed α_x and β_x parameters. For fitting a model to $\hat{\kappa}_y$ we use the definition of the small-sample-corrected AIC in Appendix C to search for the best-fitting ARIMA(p, 1, q) model with

possible outliers, where $p, q \in \{0, 1, 2, 3\}$. As described in Section 4, we only search for AO, LS or TC outliers, as we leave in IOs to avoid under-estimation of σ_ϵ^2 for financial applications. We have used a critical threshold of 3.5 to avoid over-detection of outliers, as recommended by Richards [2024, Section 4.3].

	Females		Males	
Country	Type	ω_{2020}	Type	ω_{2020}
Italy	TC	0.068026	TC	0.082307
Netherlands	LS	0.053908	LS	0.066529
Sweden	no outlier detected		TC	0.039624
UK	TC	0.060115	TC	0.058469

Table 3: Outliers found using the methodology of Chen and Liu [1993], searching for AO, LS and TC outliers with a critical threshold of 3.5. Outliers were only detected in 2020 and in no other years.

As expected, Table 3 only finds outliers (if at all) in 2020. This is partly a function of setting the critical threshold to 3.5 to avoid spurious outlier detection. Table 3 also confirms what visual inspection in Section 2 suggested: that there is no meaningful COVID-19 outlier among females in Sweden (Figure 5(b)); that males and females in the Netherlands both appear to suffer from a level shift due to COVID-19 (Figure 4(b)); and that males in Sweden and both sexes in Italy and the UK suffer from a decaying TC effect of COVID-19 (Figure 5(a), Figures 3 and 6).

However, Table 3 also produces two results that are less obvious from visual inspection: that males in the Netherlands also suffer from a level shift, despite Figure 4(a) looking like a decaying TC; and that males in Sweden suffer from a decaying TC outlier, despite Figure 5(a) looking like an AO.

We can use equation (11) to break down the components of the ARIMA forecast of κ_{y+t}. For example, Figure 9(a) shows the ARIMA forecast of κ for females in Italy, together with the contributions from the linear trend in panel (b), the TC outlier in panel (c) and the cumulative stochastic deviation in panel (d). In contrast, Figure 10(a) shows the ARIMA forecast of κ for males in the Netherlands, together with the contributions from the linear trend in panel (b), the LS outlier in panel (c) and the cumulative stochastic deviation in panel (d).

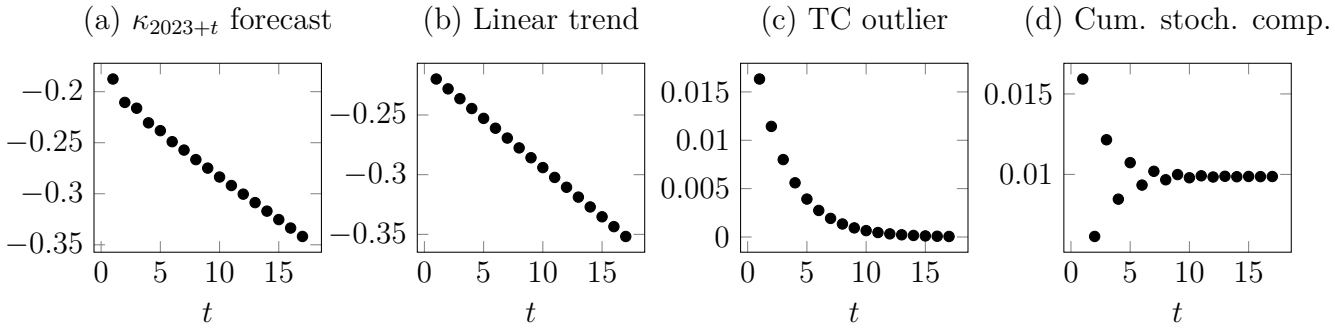

Figure 9: Illustration of a temporary change in ARIMA(1,1,0) central forecast of κ_{2023+t} for females in Italy. Panel (a) shows the forecast, while panels (b)–(d) show the components of the forecast according to equation (11).

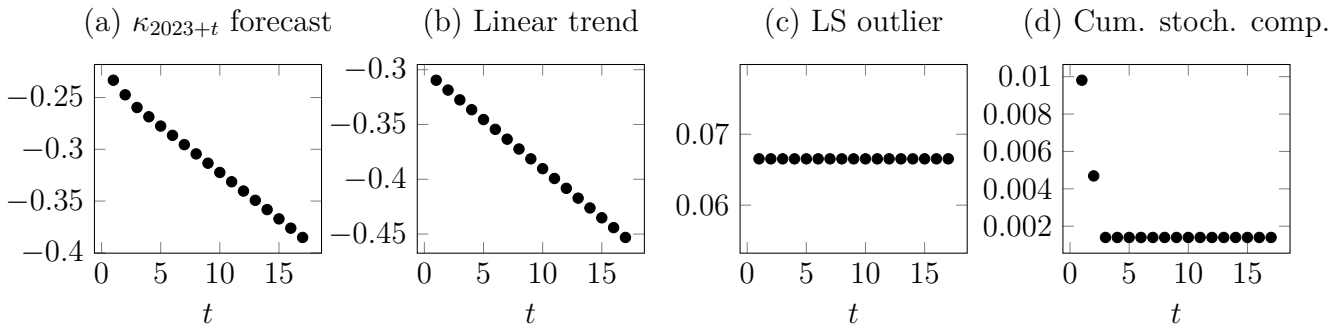

Figure 10: Illustration of a level shift outlier in ARIMA(0,1,3) central forecast of κ_{2023+t} for males in the Netherlands. Panel (a) shows the forecast, while panels (b)–(d) show the components of the forecast according to equation (11).

We have only considered outliers in ARIMA(p, 1, q) models for κ_y in a Lee-Carter model; Appendix D gives a brief overview of the relevance of outliers for mortality models that do not use ARIMA forecasts.

8 TC outliers: the true value of δ and reclassification risk

In Table 3 we found a number of TC outliers using the *a priori* value of $\delta = 0.7$ recommended by Chen and Liu [1993, Section 1.1]. We also noted in equation 12 that LS outliers are a special case of TC outliers with $\delta = 1$. This gives rise to three closely related questions: (i) whether the *a priori* value of $\delta = 0.7$ is appropriate for COVID-19 temporary changes?; (ii) whether the TC outliers in Table 3 might later be determined to be LS outliers with more data?; (iii) whether the LS outliers in Table 3 might later be determined to be TC outliers? To assess this, we plot the AICc of selected ARIMA models in Figure 11 for different values of δ.

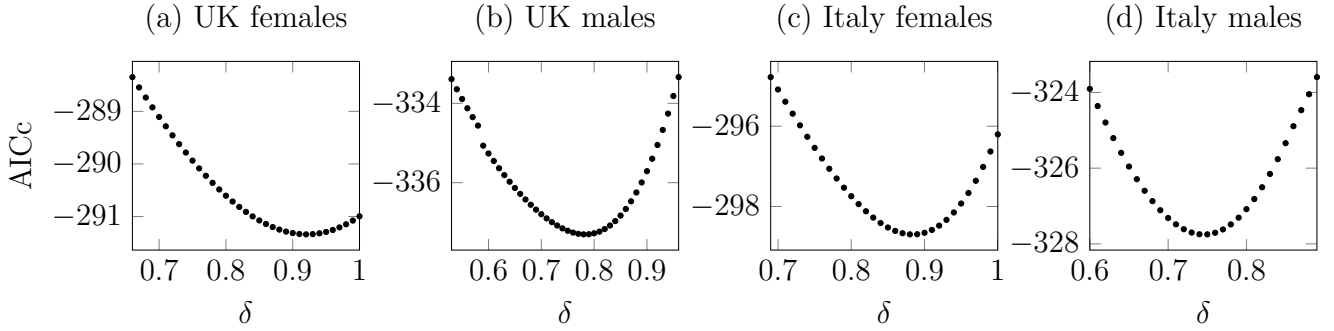

Figure 11: AICc v. δ for selected models in Table 3. The approach to calculating the AICc is described in Appendix C. The horizontal axes vary because they are chosen to include only values of δ where the AICc is within 4 units of the minimum AICc value, making the displayed range an approximate 95% confidence interval for $\hat{\delta}$.

Figure 11 shows that the optimum value of δ for TC outliers in the UK and Italy is above the default 0.7 recommended by Chen and Liu [1993, Section 1.1], although there is not yet conclusive evidence that 0.7 can be rejected. Of interest is the fact that the optimum value of δ is higher for females than for males in both Italy and the UK. In particular, the optimum value of $\hat{\delta} = 0.92$ for UK females suggests that the availability of further data beyond 2022 could yet cause the TC outlier

to be later reclassified as an LS outlier. Similarly, it is possible that the LS outliers detected for the Netherlands in Table 3 might be later reclassified as TC outliers with data beyond 2023.

With $\hat{\delta}$ for TC outliers looking likely to exceed 0.7, the effects of COVID-19 are likely to linger in mortality experience data for years to come, especially if they might be later reclassified as LS outliers. The possibility of later outlier reclassification is a particular risk for actuaries reserving for pensions and annuities, as it will likely change the ARIMA parameters and could even lead to a different choice of optimal p and q.

9 Conclusions

An ARIMA model is well suited to forecasting mortality rates dependent on a univariate time index [Kleinow and Richards, 2016]. More specifically, an ARIMA(p, 1, q) forecast can be decomposed into a linear trend and a cumulative stochastic component, as shown in equation (6). Using an ARIMA(p, 1, q) model within a Lee-Carter framework corresponds closely to a model of mortality improvements, as shown in Appendix B.

To deal with mortality patterns caused by COVID-19, we can identify various types of outlier in an ARIMA model using the methodology of Chen and Liu [1993]. For actuarial purposes we focus on handling additive outliers, level shifts and temporary changes. In contrast, we leave innovation outliers as they are, since including them in the model leads to under-statement of value-at-risk capital requirements for longevity risk [Richards, 2024, Section 4.2]. Modern mortality work often involves relatively short time series, so a critical threshold of 3.5 is used in the Chen-Liu framework to avoid spurious outlier detection.

Evidence suggests that the type of outlier caused by COVID-19 varies by country and sex. In particular, females in Sweden look like they experienced no outlier at all, while both males and females in the Netherlands experienced a level shift from 2020 onwards. Both males and females in Italy and the UK experienced a temporary change commencing in 2020 and decaying in strength thereafter. An ARIMA(p, 1, q) forecast can then be decomposed into a linear trend, an ongoing outlier effect and a cumulative stochastic component. The precise nature of the outlier caused by COVID-19 therefore needs to be acknowledged by actuaries when calibrating stochastic projection models for pricing, reserving and setting capital requirements.

Tentative evidence suggests that the value of δ for TC outliers might be higher than the default value of 0.7 recommended by Chen and Liu [1993, Section 1.1]. If so, TC outliers will have longer-lasting effects in mortality experience and forecasts. Outliers may even be later reclassified, with switches between TC and LS outliers conceivable with more experience data. This possibility warrants prudence in actuarial forecasts of mortality for pensions and annuities business, as outlier reclassification will change ARIMA forecasts.

Acknowledgments

The author thanks Professor Torsten Kleinow and Professor Angus S. Macdonald for a discussion on parameter counting with outliers, and also Professor Andrew J. G. Cairns, Gavin P. Ritchie and June Shi for comments on an earlier draft of this paper. Any errors or omissions remain the sole responsibility of the author. Calculations were performed using the Projections Toolkit and bespoke programs written in R [R Core Team, 2021]. This document was typeset in LaTeX. Graphs were created with tikz [Tantau, 2024] and pgfplots [Feuersänger, 2015].

No part of this paper, or the results contained therein, was generated using AI tools of any kind.

References

H. Akaike. Factor analysis and AIC. *Psychometrica*, 52:317–333, 1987. ISSN 0033–3123. doi: 10.1007/BF02294359.

R. W. Bacon. Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes. *Energy Economics*, 13(3):211–218, 1991. doi: 10.1016/0140-9883(91)90022-R. URL `https://www.sciencedirect.com/science/article/pii/014098839190022R`.

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

A. J. G. Cairns and D. Blake. ADM's APPLE: The Accelerated Deaths Model with an Application to the Covid-19 Pandemic. *Insurance Mathematics and Economics*, 2026. doi: 10.1016/j.insmatheco.2026.103231.

A. J. G. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73:687–718, 2006. doi: 10.1111/j.1539-6975.2006.00195.x.

A. J. G. Cairns, D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong, and I. Balevich. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1):1–35, 2009. doi: 10.1080/10920277.2009.10597538.

C. Chen and L-M. Liu. Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88(421):284–297, 1993.

I. D. Currie. Smoothing constrained generalized linear models with an application to the Lee-Carter model. *Statistical Modelling*, 13(1):69–93, 2013. doi: 10.1177/1471082X12471373.

I. D. Currie, M. Durban, and P. H. C. Eilers. Smoothing and forecasting mortality rates. *Statistical Modelling*, 4:279–298, 2004. doi: 10.1191/1471082X04st080oa.

A. Delwarde, M. Denuit, and P. H. C. Eilers. Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: a penalized likelihood approach. *Statistical Modelling*, 7:29–48, 2007. doi: 10.1177/1471082X0600700103.

V. A. B. Djeundje and I. D. Currie. Smoothing dispersed counts with applications to mortality data. *Annals of Actuarial Science*, 5(I):33–52, 2011.

C. Feuersänger. *Manual for Package PGFPLOTS, Version 1.12.1*, 2015. URL `http://sourceforge.net/projects/pgfplots`.

P. Galeano, D. Peña, and R. S. Tsay. Outlier detection in multivariate time series by projection pursuit. *Journal of the American Statistical Association*, 101(474):654–669, 2006. doi: 10.1198/0162145000001131.

G. Gardner, A. C. Harvey, and G. D. A. Phillips. Algorithm AS 154: An Algorithm for Exact Maximum Likelihood Estimation of Autoregressive-Moving Average Models by Means of Kalman Filtering. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 29(3):311–322, 1980. doi: 10.2307/2346910.

F. Girosi and G. King. *Demographic Forecasting*. Princeton University Press, 2008. ISBN 978-0-691-13095-8.

J. Goes, K. Barigou, and A. Leucht. Bayesian mortality modelling with pandemics: a vanishing jump approach. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 74:1150–1182, 2025. doi: 10.1093/jrsssc/qlaf018.

A. C. Harvey. *Time Series Models*. Philip Allan, 1981. ISBN 0-86003-032-6.

C. M. Hurvich and C. L. Tsai. Regression and time-series model selection in small samples. *Biometrika*, 76(2):297–307, 1989. doi: 10.1093/biomet/76.2.297.

IEEE Computer Society. *IEEE Standard for Floating-Point Arithmetic*. IEEE, New York, NY, USA, 2019. URL `https://ieee.org`.

J. G. Kirkby and I. D. Currie. Smooth models of mortality with period shocks. *Statistical Modelling*, 10(2):177–196, 2010. doi: 10.1017/1471082X0801000204.

T. Kleinow and S. J. Richards. Parameter risk in time-series mortality forecasts. *Scandinavian Actuarial Journal*, 2016(10):1–25, 2016. doi: 10.1080/03461238.2016.1255655.

D. E. Knuth. Two Notes on Notation. *The American Mathematical Monthly*, 99(5):403–422, 1992. doi: 10.1080/00029890.1992.11995869.

R. D. Lee and L. Carter. Modeling and forecasting US mortality. *Journal of the American Statistical Association*, 87:659–671, 1992. doi: 10.1080/01621459.1992.10475265.

A. S. Macdonald and S. J. Richards. On Contemporary Mortality Models for Actuarial Use II: Principles. *British Actuarial Journal*, 2025. doi: 10.1017/S1357321725000133.

R. D. Martin, A. Samarov, and V. J. Vandaele. Robust methods for ARIMA models. In E. Zellner, editor, *Applied Time Series Analysis of Economic Data*, pages 153–177. Washington Bureau of the Census, 1983.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL `https://www.R-project.org/`.

A. E. Renshaw and S. Haberman. On the forecasting of mortality reduction factors. *Insurance: Mathematics and Economics*, 32:379–401, 2003.

S. J. Richards. Some comments on "A Hermite-spline approach for modelling population mortality" by Tang, Li & Tickle (2022). *Annals of Actuarial Science*, 17(3):643–646, 2023. doi: 10.1017/S174849952300012X.

S. J. Richards. Robust mortality forecasting in the presence of outliers. *British Actuarial Journal*, 29:e19, 2024. doi: 10.1017/S1357321724000175.

S. J. Richards and I. D. Currie. Longevity risk and annuity pricing with the Lee-Carter model. *British Actuarial Journal*, 15(II) No. 65:317–365 (with discussion), 2009. doi: 10.1017/S1357321700005675.

S. J. Richards, I. D. Currie, T. Kleinow, and G. P. Ritchie. A stochastic implementation of the APCI model for mortality projections. *British Actuarial Journal*, 24:e13, 2019. doi: 10.1017/S1357321718000260. URL `https://www.longevitas.co.uk/apci`.

R. H. Shumway and D. S. Stoffer. *Time series analysis and its applications*. Springer, third edition, 2010. ISBN 978-1-4419-7864-6.

T. Tantau. *TikZ and PGF manual for Version 3.1.10*, 2024. URL `https://pgf-tikz.github.io/pgf/pgfmanual.pdf`.

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) — China, 2020. *China CDC Weekly*, 2(8):113–122, 2020. ISSN 2096-7071. doi: 10.46234/ccdcw2020.032. URL `http://weekly.chinacdc.cn//article/id/e53946e2-c6c4-41e9-9a9b-fea8db1a8f51`.

R. S. Tsay. *Multivariate time series analysis*. Wiley, 2014. ISBN 978-1-118-61790-8.

R. C. Willets. Mortality in the next Millennium. *Staple Inn Actuarial Society, London*, 1999.

R. C. Willets. The cohort effect: insights and explanations. *British Actuarial Journal*, 10(4):833–898, 2004. doi: 10.1017/S1357321700002762.

Appendices

A Lee-Carter identifiability constraints

The Lee-Carter model in equation (1) is not identifiable — there is an infinite number of parameterisations that produce the same fitted values $\log \hat{\mu}_{x,y}$. Various authors have therefore used different identifiability constraints. Lee and Carter [1992] used the following:

$$\sum_y \kappa_y = 0 \quad \text{and} \quad \sum_x \beta_x = 1 \tag{13}$$

while Renshaw and Haberman [2003] used the following

$$\kappa_1 = 0 \quad \text{and} \quad \sum_x \beta_x = 1 \tag{14}$$

Equation systems (13) and (14) are linear constraints that uniquely identify the parameters in the Lee-Carter model. In contrast, Girosi and King [2008, p.35] used a non-linear constraint system:

$$\sum_y \kappa_y = 0 \quad \text{and} \quad \sum_x \beta_x^2 = 1 \tag{15}$$

which does not uniquely identify the Lee-Carter parameters since $(-\beta_x)^2 = \beta_x^2$. For any set of Lee-Carter parameters satisfying equation system (15), mapping $\hat{\kappa}_y \to -\hat{\kappa}_y \ \forall y$ and $\hat{\beta}_x \to -\hat{\beta}_x \ \forall x$ will produce the same $\log \hat{\mu}_{x,y}$. Similarly, Richards and Currie [2009] used the following alternative non-linear constraint system:

$$\sum_y \kappa_y = 0 \quad \text{and} \quad \sum_y \kappa_y^2 = 1 \tag{16}$$

which, like the Girosi-King system, does not uniquely identify the Lee-Carter parameters because $(-\kappa_y)^2 = \kappa_y^2$. To make equation system (16) uniquely identify the Lee-Carter parameters we additionally specify that $\kappa_1 > \kappa_{n_y}$, where n_y is the number of years of observation. We have used the constraint system in equation (16) in this paper, together with the condition $\kappa_1 > \kappa_{n_y}$, to keep the scale of the $\hat{\kappa}_y$ estimates comparable between countries and sexes.

The choice of constraint system does not change the fitted values for $\log \mu_{x,y}$ in equation (1) apart from minor arithmetical rounding. Furthermore, while the parameter estimates $\hat{\alpha}_x$, $\hat{\beta}_x$ and $\hat{\kappa}_y$ are all affected by the choice of identifiability constraints, forecasts of $\log \mu_{x,y}$ using an ARIMA model for κ_y are invariant to the choice of identifiability constraint for the Lee-Carter and APC models. This is not the case for all forecasting models — for example, the APCI model [Richards et al., 2019] requires five identifiability constraints, and forecasts of $\log \mu_{x,y}$ under the APCI model are not always invariant to the choice of those identifiability constraints.

B Mortality improvements and ARIMA forecasting of κ_y

Let $q_{x,y}$ be the one-year probability of death aged x in year y. Willets [1999, 2004] introduced the concept of a mortality improvement rate defined as follows:

$$1 - \frac{q_{x,y}}{q_{x,y-1}} \tag{17}$$

Formula (17) produces positive values when mortality rates are falling. We can also use formula (17) for the mortality hazard and, in doing so, we note an aspect of the Lee-Carter model in equation (1):

$$1 - \frac{\mu_{x,y+t}}{\mu_{x,y+t-1}} = 1 - \frac{\exp(\alpha_x + \beta_x \kappa_{y+t})}{\exp(\alpha_x + \beta_x \kappa_{y+t-1})} \tag{18}$$

$$= 1 - \exp(\beta_x(\kappa_{y+t} - \kappa_{y+t-1})) \tag{19}$$

$$= 1 - \exp(\beta_x \Delta\kappa_{y+t}) \tag{20}$$

$$= 1 - \left[1 + \beta_x \Delta\kappa_{y+t} + \frac{1}{2}(\beta_x \Delta\kappa_{y+t})^2 + \dots\right] \tag{21}$$

$$\approx -\beta_x \Delta\kappa_{y+t} \tag{22}$$

The justification for the last step is that, even including the outliers, the largest absolute value of $\Delta\hat{\kappa}_{y+t}$ in Figures 3-6 is around 0.06. The largest absolute value of $\hat{\beta}_x$ (not shown) is around 3, so $\frac{1}{2}(\hat{\beta}_x \Delta\hat{\kappa}_{y+t})^2$ is typically much smaller than 0.0162, which is less than a tenth of the equivalent $\hat{\beta}_x \Delta\hat{\kappa}_{y+t}$ value. Thus, observed mortality improvements are broadly equivalent to $\Delta\hat{\kappa}_{y+t}$ after appropriate scaling for age under the Lee-Carter model. Since an ARIMA(p, 1, q) model for κ_{y+t} is a ARMA(p, q) model for $\Delta\kappa_{y+t}$, there is a close link between forecasting mortality improvements and forecasting κ_{y+t} using an ARIMA(p, 1, q) model.

C ARIMA model selection with outliers

When choosing the ARIMA model that best fits the estimated $\hat{\kappa}_{y+t}$ series, a useful tool is an information criterion such as the AIC from Akaike [1987]. This allows for the fact that different ARIMA models have different numbers of parameters. However, there are two aspects to bear in mind for mortality work. The first is that the sequence length — the number of $\hat{\kappa}_{y+t}$ estimates — is typically rather short at around $n_y = 50$ observations [Richards, 2024, Appendix C]. As such, it is a good idea to use the definition of the AIC with a small-sample correction when choosing an ARIMA model for κ_{y+t}. We use the definition of Hurvich and Tsai [1989] for the AICc as follows:

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{n_y - k - 1} \tag{23}$$

where AIC is the usual definition of Akaike's Information Criterion and k is the number of parameters estimated in the ARIMA model. The denominator in equation 23 means that we reject any ARIMA model that has $k \geq n_y - 1$ parameters.

The second aspect lies with the parameter count, k. In an ARIMA$(p, 1, q)$ model without outliers there are typically $k = 2 + p + q$ parameters: (i) a variance estimate, $\hat{\sigma}^2_\epsilon$, (ii) a drift term, $\hat{\mu}$, (iii) p estimates of autoregressive coefficients and (iv) q estimates of moving-average coefficients. However, an ARIMA model with m co-estimated outlier parameters actually has $k = 2+p+q+3m$ parameters. It is necessary to make this allowance, as it is possible for an ARIMA$(p_1, 1, q_1)$ model to be estimated without finding any outliers (i.e. $m = 0$), whereas an ARIMA$(p_2, 1, q_2)$ model fitted to the same data might find $m > 0$ outliers. Since any outlier parameters are co-estimated along with the ARIMA parameters, the total number of outlier parameters needs to be included in the overall parameter count for consistency. The multiplier 3 stems from the fact that each outlier has (i) a type, (ii) a location and (iii) an effect. A consequence of this in the denominator of equation 23 is that we reject any ARIMA model that 'identifies' $m \geq (n_y - p - q)/3 - 1$ outliers.

Note that the multiplier 3 does not change if we set an *a priori* value for δ for a TC or estimate $\hat{\delta}$ from the data. As shown in equation (12), the only thing distinguishing AO, LS and TC outliers

is the value of δ. It makes no difference to the parameter count if the domain of δ is restricted to a small set like $\{0, 0.7, 1\}$ or is free to take values in the real interval $[0, 1]$.

D Other stochastic projection models

The main body of the paper considers only mortality models that rely on an ARIMA time-series forecast of a univariate index, κ_y. However, there are many important forecasting models that do not. Examples include penalty projections [Currie et al., 2004], bivariate random walks with drift [Cairns et al., 2006] and trivariate random walks with drift [Cairns et al., 2009]. Four-dimensional random walks are possible in theory, but in practice the data often better support the replacement of one of the time indices with a constant, thus reducing the forecast to a trivariate random walk [Richards, 2023].

D.1 2DAP model

Currie et al. [2004] presented a two-dimensional, penalised-spline model for mortality in age and time (the 2DAP model). Forecasting is done by extrapolating the penalty function in the time dimension. Kirkby and Currie [2010] presented an extension of the 2DAP model to include period effects like mortality shocks. There is no attempt to identify outliers or their location, as period effects are estimated for every year. Since there is no attempt to locate outliers, there is also no attempt to classify their type. Note that while period effects are regarded as independent of each other, there is the option for an indirect link in Kirkby and Currie [2010] — a smoothing constant is applied to each period effect, and this smoothing constant can be scaled relative to the size of the largest period effect. This results in heavier smoothing for minor period effects; see Richards [2024, Figures 11 and 12] for an illustration.

The net effect of the model from Kirkby and Currie [2010] is to make the smoothing splines in calendar time more closely identified with the local trend in mortality, with annual deviations from the trend accounted for by the period effects. Note that what counts as 'local trend' depends on the knot spacing used for the splines (a knot spacing of five or more years is typically used in the time dimension). Any potential distortion from outliers is absorbed into the period effects, thus making the penalty projection a more stable extrapolation of recent trend mortality; see Richards [2024, Figure 13] for an illustration. Thus, the 2DAP model with period effects does not need to classify the type of outlier.

D.2 Multivariate random walks with drift

The Cairns-Blake-Dowd and Tang-Lee-Tickle model families forecast mortality via a multivariate random walk with drift. Richards [2024] allowed for outliers using the detection methodology of Galeano et al. [2006], which identifies the location of outliers, but does not classify them. This is because a random walk with drift is memoryless — the next value in time depends only on (i) the current value, (ii) the drift term, and (iii) a random innovation uncorrelated with all other innovation terms. There is therefore no scope for identifying the outlier type, as there is for an ARIMA model for a univariate index. This would change if the multivariate random walk were replaced with, say, a more general vector ARMA (VARMA) model [Tsay, 2014] for differences in the time index. However, reliably estimating the many parameters of a VARMA model would be challenging for mortality forecasting, as typically there are only fifty or so relevant observations in a sequence.

An important aspect of the application of Galeano et al. [2006] to mortality forecasting is that there are fewer false positives when using differenced multivariate mortality series, as opposed to using the undifferenced series [Richards, 2024, Appendix C].

Contact

More information including case studies, latest features, technical documentation and demonstration videos can be found on our website at www.longevitas.co.uk

LONGEVITAS™